

CSI REPORTING

TECHNICAL FIELD

[0001] The present disclosure relates to wireless communications, and more specifically to user equipments (UEs), base stations, processors, and methods for channel state information (CSI) reporting.

BACKGROUND

[0002] A wireless communications system may include one or multiple network communication devices, such as base stations, which may be otherwise known as an eNodeB (eNB), a next-generation NodeB (gNB), or other suitable terminology. Each network communication devices, such as a base station may support wireless communications for one or multiple user communication devices, which may be otherwise known as user equipment (UE), or other suitable terminology. The wireless communications system may support wireless communications with one or multiple user communication devices by utilizing resources of the wireless communication system (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers). Additionally, the wireless communications system may support wireless communications across various radio access technologies including third generation (3G) radio access technology, fourth generation (4G) radio access technology, fifth generation (5G) radio access technology, among other suitable radio access technologies beyond 5G (e.g., sixth generation (6G)).

[0003] Some artificial intelligence or machine learning (AI/ML) models based CSI functionalities (such as AI/ML based CSI compression, AI/ML based CSI prediction, and the like) have been introduced, they may provide performance gain in most use cases. However, there are still some open problems related to support of the AI/ML ability in these functionalities that need to be studied in the future.

SUMMARY

[0004] The present disclosure relates to methods, apparatuses, and systems that support CSI reporting.

[0005] In a first aspect of the solution, a user equipment (UE) may comprise: a processor; and a transceiver coupled to the processor, wherein the processor is configured

to: transmit, via the transceiver and to a base station, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and receive, via the transceiver and from the base station, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0006] In some implementations of the method and apparatuses described herein, the one or more AI/ML functionalities may include at least one of: an AI/ML functionality for CSI compression; an AI/ML functionality for CSI prediction; an AI/ML functionality for spatial domain beam prediction, and an AI/ML functionality for temporal domain beam prediction.

[0007] In some implementations of the method and apparatuses described herein, the capability information may include at least one of a first capability information on the AI/ML functionality for CSI compression and a second capability information on the AI/ML functionality for CSI prediction.

[0008] In some implementations of the method and apparatuses described herein, the first capability information may include: a first information on CSI compression reporting type, wherein the CSI compression reporting type includes at least one of per-rank CSI reporting and per-layer CSI reporting; and a second information on a number of uplink control information (UCI) bits for a compressed CSI report; and a third information on whether a vector-based quantization scheme or a scalar based quantization scheme for compressed CSI feedback is supported by the UE.

[0009] In some implementations of the method and apparatuses described herein, the second capability information may include: a fourth information on CSI type(s) to be predicted, wherein the CSI type(s) to be predicted includes at least one of rank indicator (RI), precoding matrix indicator (PMI), layer indicator (LI), and channel quality indicator (CQI); and a fifth information on a relation between a number of historical CSI-reference signal (RS) receptions within a measurement time window and a number of predicted CSIs in a prediction time window.

[0010] In some implementations of the method and apparatuses described herein, in the case that the CSI type(s) to be predicted includes at least one of the PMI and the CQI,

the second capability information may further include: a sixth information on whether at least one of wideband CSI prediction and subband CSI prediction is supported by the UE; and a seventh information on a codebook type for PMI supported by the UE.

[0011] In some implementations of the method and apparatuses described herein, each of the at least one AI/ML CSI report configuration may be transmitted via a preset higher layer information element configured by the base station.

[0012] In some implementations of the method and apparatuses described herein, an AI/ML CSI report configuration among the at least one AI/ML CSI report configuration may include at least one of: a configuration identifier (ID) indicating the AI/ML CSI report configuration identity; a resource information on CSI resource for channel measurement or CSI resource for interference measurement; an AI/ML functionality or model information indicating an AI/ML functionality or model associated with the AI/ML CSI report configuration; a usage information indicating the usage of the AI/ML CSI report configuration; a reporting type for the AI/ML CSI report configuration, wherein the reporting type includes one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic; a rank subset restriction indicating an allowed rank indicator (RI) report for the AI/ML CSI report configuration; and an associated CSI report configuration for fallback operation, which indicates a non-AI/ML CSI report configuration for fallback operation without AI/ML operation.

[0013] In some implementations of the method and apparatuses described herein, in the case that the UE determines to perform a fallback operation without AI/ML operation based on the AI/ML CSI report configuration, the processor may be configured to: drop an AI/ML CSI report corresponding to the AI/ML CSI report configuration, and transmit, via the transceiver and to the base station, a non-AI/ML CSI report corresponding to the associated non-AI/ML CSI report configuration, wherein the same CSI resources for channel and/or interference measurement for the AI/ML CSI report configuration is used for the associated non-AI/ML CSI report configuration.

[0014] In some implementations of the method and apparatuses described herein, in case that there are more than one AI/ML functionality identified at the UE, the AI/ML functionality or model information may include a functionality ID of the AI/ML functionality associated with the AI/ML CSI report configuration; and in case that there

is only one AI/ML functionality identified at the UE, the AI/ML functionality or model information may include an AI/ML model ID for the AI/ML functionality associated with the AI/ML CSI report configuration, or the AI/ML functionality or model information may indicate to enable the AI/ML functionality associated with the AI/ML CSI report configuration.

[0015] In some implementations of the method and apparatuses described herein, the usage of the AI/ML CSI report configuration may include at least one of: data collection for model training, model inference, and performance monitoring, wherein the model inference includes at least one of model inference for CSI compression, model inference for CSI prediction, model inference for spatial domain beam prediction, and model inference for temporal domain beam prediction.

[0016] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for data collection for model training for two-sided model for CSI compression, a first CSI obtained by using an AI/ML model at the UE together with a second CSI obtained without using any AI/ML model at the UE, which is used as a model input for the AI/ML model at the UE, may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0017] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for inference for CSI compression, and a number of uplink control information (UCI) bits for a compressed CSI report is determined by the UE, the number of UCI bits for the compressed CSI report may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and a compressed CSI may be included in a second part of the AI/ML CSI report.

[0018] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for inference for CSI prediction, a number of time instance of a predicted CSI and a predicted CSI corresponding to a first time instance may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and other predicted CSI corresponding to other time instance may be included in a second part of the AI/ML CSI report.

[0019] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for performance monitoring, a performance gap between a reported CSI and the actual CSI of a same CSI-RS resource on a same transmission occasion may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0020] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for performance monitoring, a ground-truth CSI corresponding to a CSI-RS resource or a CSI-RS resource set with a certain quantization scheme may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0021] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may further include at least one non-AI/ML CSI report configuration.

[0022] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration includes at least one of: at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink control channel (PUCCH), and at least one non-AI/ML CSI report configuration for SP CSI report on PUCCH, and the processor may be further configured to: receive, via the transceiver and from the base station, a media access control (MAC) control element (CE), wherein the MAC CE includes a AI/ML field indicating whether the MAC CE is used for activating at least one AI/ML CSI report configuration for SP CSI report on PUCCH or at least one non-AI/ML CSI report configuration for SP CSI report on PUCCH.

[0023] In some implementations of the method and apparatuses described herein, the AI/ML field may include one bit indicating whether a predetermined field in the MAC CE is used for AI/ML or non-AI/ML, in the case that the one bit has a first value, each bit in the predetermined field may indicate whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated, and in the case that the one bit has a second value different from the first value, each bit in the predetermined field may indicate whether a corresponding non-AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated.

[0024] In some implementations of the method and apparatuses described herein, the AI/ML field may include a plurality of bits, each of which indicates whether a

corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated.

[0025] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may include at least one of: at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink shared channel (PUSCH), and at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH.

[0026] In some implementations of the method and apparatuses described herein, the processor may be further configured to: receive, via the transceiver and from the base station, a SP CSI report trigger state indication, each CSI report trigger state indicated in the SP CSI report trigger state indication is mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH, wherein a CSI report trigger state indicated in the SP CSI report trigger state indication may be associated with a configuration identifier (ID) of an AI/ML CSI report configuration for SP CSI report on PUSCH.

[0027] In some implementations of the method and apparatuses described herein, the processor may be further configured to: receive, via the transceiver and from the base station, a first SP CSI report trigger state indication and a second SP CSI report trigger state indication, each CSI report trigger state indicated in the first SP trigger state indication and the second SP CSI report trigger state indication is mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH, wherein the first SP CSI report trigger state indication is associated with at least one AI/ML CSI report configuration for SP CSI report on PUSCH, and the second SP CSI report trigger state indication is associated with at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH.

[0028] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may include at least one of: at least one AI/ML CSI report configuration for aperiodic (AP) CSI report, and at least one non-AI/ML CSI report configuration for AP CSI report.

[0029] In some implementations of the method and apparatuses described herein, the processor may be further configured to: receive, via the transceiver and from the base station, a AP CSI report trigger state indication, each CSI report trigger state indicated in

the AP CSI report trigger state indication is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report, wherein a CSI report trigger state indicated in the AP CSI report trigger state indication is associated with a configuration ID of an AI/ML CSI report configuration for AP CSI report.

[0030] In some implementations of the method and apparatuses described herein, the processor may be further configured to: receive, via the transceiver and from the base station, a first AP CSI report trigger state indication and a second AP CSI report trigger state indication, each CSI report trigger state indicated in the first AP CSI report trigger state indication and the second AP CSI report trigger state indication is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report, wherein the first AP CSI report trigger state indication is associated with at least one AI/ML CSI report configuration for AP CSI report, and the second AP CSI report trigger state indication is associated with at least one non-AI/ML CSI report configuration for AP CSI report.

[0031] In some implementations of the method and apparatuses described herein, the processor may be further configured to: determine priorities among AI/ML CSI reports based on usage of the AI/ML CSI reports.

[0032] In some implementations of the method and apparatuses described herein, a priority of an AI/ML CSI report for performance monitoring may be higher than a priority of an AI/ML CSI report for inference, and the priority of the AI/ML CSI report for inference may be higher than a priority of an AI/ML CSI report for data collection for model training.

[0033] In some implementations of the method and apparatuses described herein, the processor may be further configured to: determine priorities of CSI reports at least based on the at least one CSI report configuration.

[0034] In some implementations of the method and apparatuses described herein, the priorities of the CSI reports may be determined based on priority values of the CSI reports, wherein a priority value of a CSI report may be determined based on at least one of: a type indication value indicating a reporting type of a CSI report configuration corresponding to the CSI report, wherein the reporting type includes one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic; a configuration identifier (ID) of the CSI report

configuration corresponding to the CSI report; a measurement indication value indicating whether the CSI report carries L1-reference signal receiving power (RSRP) or L1-signal to interference plus noise ratio (SINR) or not; the max number of each type of CSI report configurations in the at least one CSI report configuration, wherein types of CSI report configuration at least include AI/ML type and non-AI/ML type; a serving cell index for the UE; and the max number of serving cells.

[0035] In some implementations of the method and apparatuses described herein, a priority of a first CSI report may be higher than a priority of a second CSI report in the case that a priority value of the first CSI report is lower than a priority value of the second CSI report.

[0036] In some implementations of the method and apparatuses described herein, the priority value of the CSI report is determined by calculating a sum of: a value equals to twice the product of a sum value of the max number of each type of the CSI report configurations in the at least one CSI report configuration, the max number of serving cells, and the type indication value for the CSI report; a product of the sum value, the max number of serving cells, and the measurement indication value for the CSI report; a product of the sum value and the serving cell index for the CSI report; and a value related to the configuration ID for the CSI report, wherein the value related to the configuration ID for the CSI report is configured to make each AI/ML CSI report has a higher priority value than any non-AI/ML CSI report.

[0037] In some implementations of the method and apparatuses described herein, the priority value of the CSI report may be determined by calculating a sum of: a value equals to twice the product of the max number of first CSI report configurations in the at least one CSI report configuration, the max number of serving cells, and the type indication value for the CSI report, wherein the first CSI report configurations have the same type to a CSI report configuration corresponding to the CSI report; a product of the max number of the first CSI report configurations, the max number of serving cells, and the measurement indication value for the CSI report; a product of the sum value and the serving cell index for the CSI report; and the configuration ID for the CSI report.

[0038] In some implementations of the method and apparatuses described herein, in the case that time occupancy of physical uplink shared channel (PUSCH) or physical uplink control channel (PUCCH) to carry CSI reports overlap in at least one orthogonal

frequency-division multiplexing (OFDM) symbol and the CSI reports are transmitted on the same carrier, the processor may be further configured to: determine reporting priorities of the CSI reports based on a predetermined rule; and drop one of the CSI reports based on the reporting priorities.

[0039] In some implementations of the method and apparatuses described herein, the predetermined rule may include: an AI/ML CSI report has a lower reporting priority than a non-AI/ML CSI report.

[0040] In some implementations of the method and apparatuses described herein, the predetermined rule may include: allocating reporting priorities, in a priority order from high to low, to an AP non-AI/ML CSI report, an AP AI/ML CSI report, a SP non-AI/ML CSI report on PUSCH, a SP AI/ML CSI report on PUSCH, a SP non-AI/ML CSI report on PUCCH, a SP AI/ML CSI report on PUCCH, a periodic non-AI/ML CSI report, and a periodic AI/ML CSI report.

[0041] In some implementations of the method and apparatuses described herein, the predetermined rule may include: allocating reporting priorities, in a priority order from high to low, to an AP non-AI/ML CSI report, a SP non-AI/ML CSI report on PUSCH, a SP non-AI/ML CSI report on PUCCH, a periodic non-AI/ML CSI report, an AP AI/ML CSI report, a SP AI/ML CSI report on PUSCH, a SP AI/ML CSI report on PUCCH, and a periodic AI/ML CSI report.

[0042] In a second aspect of the solution, a base station may comprise: a processor; and a transceiver coupled to the processor, wherein the processor is configured to: receive, via the transceiver and from a user equipment (UE), capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and transmit, via the transceiver and to the UE, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0043] In some implementations of the method and apparatuses described herein, the one or more AI/ML functionalities may include at least one of: an AI/ML functionality for CSI compression; an AI/ML functionality for CSI prediction; an AI/ML functionality for spatial domain beam prediction, and an AI/ML functionality for temporal domain beam prediction.

[0044] In some implementations of the method and apparatuses described herein, the capability information may include at least one of a first capability information on the AI/ML functionality for CSI compression and a second capability information on the AI/ML functionality for CSI prediction.

[0045] In some implementations of the method and apparatuses described herein, the first capability information may include: a first information on CSI compression reporting type, wherein the CSI compression reporting type includes at least one of per-rank CSI reporting and per-layer CSI reporting; and a second information on a number of uplink control information (UCI) bits for a compressed CSI report; and a third information on whether a vector-based quantization scheme or a scalar based quantization scheme for compressed CSI feedback is supported by the UE.

[0046] In some implementations of the method and apparatuses described herein, the second capability information may include: a fourth information on CSI type(s) to be predicted, wherein the CSI type(s) to be predicted includes at least one of rank indicator (RI), precoding matrix indicator (PMI), layer indicator (LI), and channel quality indicator (CQI); and a fifth information on a relation between a number of historical CSI-reference signal (RS) receptions within a measurement time window and a number of predicted CSIs in a prediction time window.

[0047] In some implementations of the method and apparatuses described herein, in the case that the CSI type(s) to be predicted includes at least one of the PMI and the CQI, the second capability information may further include: a sixth information on whether at least one of wideband CSI prediction and subband CSI prediction is supported by the UE; and a seventh information on a codebook type for PMI supported by the UE.

[0048] In some implementations of the method and apparatuses described herein, each of the at least one AI/ML CSI report configuration may be transmitted via a preset higher layer information element configured by the base station.

[0049] In some implementations of the method and apparatuses described herein, an AI/ML CSI report configuration among the at least one AI/ML CSI report configuration may include at least one of: a configuration identifier (ID) indicating the AI/ML CSI report configuration identity; a resource information on CSI resource for channel measurement or CSI resource for interference measurement; an AI/ML functionality or model information indicating a AI/ML functionality or model associated with the AI/ML

CSI report configuration; a usage information indicating the usage of the AI/ML CSI report configuration; a reporting type for the AI/ML CSI report configuration, wherein the reporting type includes one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic; a rank subset restriction indicating an allowed rank indicator (RI) report for the AI/ML CSI report configuration; and a CSI report configuration indication for fallback operation, which indicates a non-AI/ML CSI report configuration for fallback operation without AI/ML operation.

[0050] In some implementations of the method and apparatuses described herein, in case that there are more than one AI/ML functionality identified at the UE, the AI/ML functionality or model information may include a functionality ID of the AI/ML functionality associated with the AI/ML CSI report configuration; and in case that there is one AI/ML functionality identified at the UE, the AI/ML functionality or model information may include an AI/ML model ID for the AI/ML functionality associated with the AI/ML CSI report configuration, or the AI/ML functionality or model information may indicate to enable the AI/ML functionality associated with the AI/ML CSI report configuration.

[0051] In some implementations of the method and apparatuses described herein, the usage of the AI/ML CSI report configuration may include at least one of: data collection for model training, model inference, and performance monitoring, wherein the model inference includes at least one of model inference for CSI compression, model inference for CSI prediction, model inference for spatial domain beam prediction, and model inference for temporal domain beam prediction.

[0052] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for data collection for model training for two-sided model for CSI compression, a first CSI obtained by using an AI/ML model at the UE together with a second CSI obtained without using any AI/ML model at the UE, which is used as a model input for the AI/ML model at the UE, may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0053] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for inference for CSI compression, and a number of uplink control information (UCI) bits for a compressed

CSI report is determined by the UE, the number of UCI bits for the compressed CSI report may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and a compressed CSI may be included in a second part of the AI/ML CSI report.

[0054] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for inference for CSI prediction a number of time instance of a predicted CSI and a predicted CSI corresponding to a first time instance may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and other predicted CSI corresponding to other time instance may be included in a second part of the AI/ML CSI report.

[0055] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for performance monitoring, a performance gap between a reported CSI and an actual CSI of a same CSI-RS resource on a same transmission occasion may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0056] In some implementations of the method and apparatuses described herein, in the case that the AI/ML CSI report configuration is configured for performance monitoring a ground-truth CSI corresponding to a CSI-RS resource or a CSI-RS resource set with a certain quantization scheme may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

[0057] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may further include at least one non-AI/ML CSI report configuration.

[0058] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may include at least one of: at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink control channel (PUCCH), and at least one non-AI/ML CSI report configuration for SP CSI report on PUCCH, and the processor may be further configured to: transmit, via the transceiver and to the UE, a media access control (MAC) control element (CE), wherein the MAC CE includes a AI/ML field indicating whether the MAC CE is used for

activating at least one AI/ML CSI report configuration for SP CSI report on PUCCH or at least one non-AI/ML CSI report configuration for SP CSI report on PUCCH.

[0059] In some implementations of the method and apparatuses described herein, the AI/ML field may include one bit indicating whether a predetermined field in the MAC CE is used for AI/ML or non-AI/ML, in the case that the one bit has a first value, each bit in the predetermined field may indicate whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated, and in the case that the one bit has a second value different from the first value, each bit in the predetermined field may indicate whether a corresponding non-AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated.

[0060] In some implementations of the method and apparatuses described herein, the AI/ML field may include a plurality of bits, each of which indicates whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated.

[0061] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration includes at least one of: at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink shared channel (PUSCH), and at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH.

[0062] In some implementations of the method and apparatuses described herein, the processor may be further configured to: transmit, via the transceiver and to the UE, a SP CSI report trigger state indication, each CSI report trigger state indicated in the SP CSI report trigger state indication is mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH, wherein a CSI report trigger state indicated in the SP CSI report trigger state indication is associated with a configuration identifier (ID) of an AI/ML CSI report configuration for SP CSI report on PUSCH.

[0063] In some implementations of the method and apparatuses described herein, the processor may be further configured to: transmit, via the transceiver and to the UE, a first SP CSI report trigger state indication and a second SP CSI report trigger state indication, each CSI report trigger state indicated in the first SP CSI report trigger state indication and the second SP CSI report trigger state indication is mapped to a CSI request field

codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH, wherein the first SP CSI report trigger state indication is associated with at least one AI/ML CSI report configuration for SP CSI report on PUSCH, and the second SP CSI report trigger state indication is associated with at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH.

[0064] In some implementations of the method and apparatuses described herein, the at least one CSI report configuration may include at least one of: at least one AI/ML CSI report configuration for aperiodic (AP) CSI report, and at least one non-AI/ML CSI report configuration for AP CSI report.

[0065] In some implementations of the method and apparatuses described herein, the processor may be further configured to: transmit, via the transceiver and to the UE, a AP CSI report trigger state indication, each CSI report trigger state indicated in the AP CSI report trigger state indication is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report, wherein a CSI report trigger state in the AP CSI report trigger state indication may be associated with a configuration ID of an AI/ML CSI report configuration for AP CSI report.

[0066] In some implementations of the method and apparatuses described herein, the processor may be further configured to: transmit, via the transceiver and to the UE, a first AP CSI report trigger state indication and a second AP CSI report trigger state indication, each CSI report trigger state indicated in the first AP CSI report trigger state indication and the second AP CSI report trigger state indication is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report, wherein the first AP CSI report trigger state indication is associated with at least one AI/ML CSI report configuration for AP CSI report, and the second AP CSI report trigger state indication is associated with at least one non-AI/ML CSI report configuration for AP CSI report.

[0067] In a first aspect of the solution, a processor for wireless communication may comprise: at least one memory; and controller coupled with the at least one memory and configured to cause the processor to: transmit, via a transceiver and to a base station, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and receive, via the transceiver and from the base station, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated

with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0068] In a fourth aspect of the solution, a method performed by a user equipment (UE), the method may comprise: transmitting, via a transceiver and to a base station, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and receiving, via the transceiver and from the base station, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0069] In a fifth aspect of the solution, a processor for wireless communication may comprise: at least one memory; and a controller coupled with the at least one memory and configured to cause the processor to: receive, via a transceiver and from a user equipment (UE), capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and transmit, via the transceiver and to the UE, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0070] In a sixth aspect of the solution, a method performed by a base station, the method may comprise: receiving, via a transceiver and from a user equipment (UE), capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and transmitting, via the transceiver and to the UE, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0071] It is to be understood that the summary section is not intended to identify key or essential features of embodiments of the present disclosure, nor is it intended to be used to limit the scope of the present disclosure. Other features of the present disclosure will become easily comprehensible through the following description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0072] FIG. 1 illustrates an example of a wireless communications system that supports CSI reporting in accordance with aspects of the present disclosure.

[0073] FIG. 2 illustrates an example of signalling procedure for CSI reporting in accordance with aspects of the present disclosure.

[0074] FIGS. 3A to 3C illustrate examples of semi-persistent(SP) AI/ML CSI report on physical uplink control channel (PUCCH) activation/deactivation media access control (MAC) control element (CE) in accordance with aspects of the present disclosure.

[0075] FIGS. 4 and 5 illustrate examples of devices that support CSI reporting in accordance with aspects of the present disclosure.

[0076] FIGS. 6 and 7 illustrate examples of processors that support CSI reporting in accordance with aspects of the present disclosure.

[0077] FIGS. 8 and 9 illustrate flowcharts of methods that support CSI reporting in accordance with aspects of the present disclosure.

DETAILED DESCRIPTION

[0078] Principles of the present disclosure will now be described with reference to some embodiments. It is to be understood that these embodiments are described only for the purpose of illustration and help those skilled in the art to understand and implement the present disclosure, without suggesting any limitation as to the scope of the disclosure. The disclosure described herein may be implemented in various manners other than the ones described below.

[0079] In the following description and claims, unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skills in the art to which this disclosure belongs.

[0080] References in the present disclosure to “one embodiment,” “an example embodiment,” “an embodiment,” “some embodiments,” and the like indicate that the embodiment(s) described may include a particular feature, structure, or characteristic, but it is not necessary that every embodiment includes the particular feature, structure, or characteristic. Moreover, such phrases do not necessarily refer to the same embodiment(s). Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the

art to affect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

[0081] It shall be understood that although the terms “first” and “second” or the like may be used herein to describe various elements, these elements should not be limited by these terms. These terms are only used to distinguish one element from another element. For example, a first element could also be termed as a second element, and similarly, a second element could also be termed as a first element, without departing from the scope of embodiments. As used herein, the term “and/or” includes any and all combinations of one or more of the listed terms.

[0082] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of example embodiments. As used herein, the singular forms “a”, “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprises”, “comprising”, “has”, “having”, “includes” and/or “including”, when used herein, specify the presence of stated features, elements, and/or components etc., but do not preclude the presence or addition of one or more other features, elements, components and/ or combinations thereof.

[0083] As used herein, the term “communication network” refers to a network following any suitable communication standards, such as, 5G new radio (NR), long term evolution (LTE), LTE-advanced (LTE-A), wideband code division multiple access (WCDMA), high-speed packet access (HSPA), narrow band internet of things (NB-IoT), and so on. Further, the communications between a terminal device and a network device in the communication network may be performed according to any suitable generation communication protocols, including but not limited to, the first generation (1G), the second generation (2G), 2.5G, 2.75G, the third generation (3G), the fourth generation (4G), 4.5G, the fifth generation (5G) communication protocols, and/or any other protocols either currently known or to be developed in the future. Embodiments of the present disclosure may be applied in various communication systems. Given the rapid development in communications, there will also be future type communication technologies and systems in which the present disclosure may be embodied. It should not be seen as limiting the scope of the present disclosure to only the aforementioned systems.

[0084] As used herein, the term “network device” generally refers to a node in a communication network via which a terminal device can access the communication network and receive services therefrom. The network device may refer to a base station (BS) or an access point (AP), for example, a node B (NodeB or NB), a radio access network (RAN) node, an evolved NodeB (eNodeB or eNB), a NR NB (also referred to as a gNB), a remote radio unit (RRU), a radio header (RH), an infrastructure device for a V2X (vehicle-to-everything) communication, a transmission and reception point (TRP), a reception point (RP), a remote radio head (RRH), a relay, an integrated access and backhaul (IAB) node, a low power node such as a femto BS, a pico BS, and so forth, depending on the applied terminology and technology.

[0085] As used herein, the term “terminal device” generally refers to any end device that may be capable of wireless communications. By way of example rather than a limitation, a terminal device may also be referred to as a communication device, a user equipment (UE), an end user device, a subscriber station (SS), an unmanned aerial vehicle (UAV), a portable subscriber station, a mobile station (MS), or an access terminal (AT). The terminal device may include, but is not limited to, a mobile phone, a cellular phone, a smart phone, a voice over IP (VoIP) phone, a wireless local loop phone, a tablet, a wearable terminal device, a personal digital assistant (PDA), a portable computer, a desktop computer, an image capture terminal device such as a digital camera, a gaming terminal device, a music storage and playback appliance, a vehicle-mounted wireless terminal device, a wireless endpoint, a mobile station, laptop-embedded equipment (LEE), laptop-mounted equipment (LME), a USB dongle, a smart device, wireless customer-premises equipment (CPE), an internet of things (IoT) device, a watch or other wearable, a head-mounted display (HMD), a vehicle, a drone, a medical device (for example, a remote surgery device), an industrial device (for example, a robot and/or other wireless devices operating in an industrial and/or an automated processing chain contexts), a consumer electronics device, a device operating on commercial and/or industrial wireless networks, and the like. In the following description, the terms: “terminal device,” “communication device,” “terminal,” “user equipment” and “UE,” may be used interchangeably.

[0086] CSI compression based CSI report using two-sided model (such as artificial intelligence (AI)/machine learning (ML) models) has been evaluated in Release 18, where eigenvectors of downlink (DL) channel matrix were firstly compressed by the UE via a

UE side CSI generation model to obtain a compressed CSI, which was reported to the NW by a CSI reporting procedure. The NW reconstructs the eigenvectors of the DL channel matrix by inputting the received compressed CSI to a NW side CSI reconstruction model. CSI compression based CSI report using two-sided model may provide performance gain in most use cases. Considering such benefit, the two-sided AI/ML model may be further specified for CSI report enhancement in Release 19. One method to enable the AI/ML based CSI compression is to enhance the legacy CSI reporting framework by configuring necessary AI/ML related parameters for each CSI report configuration and specify the corresponding UE behavior.

[0087] CSI framework has been introduced in NR release 15 and specified in 3GPP Technical Specification TS38.214 V15.0.0. For example, procedures on aperiodic CSI reporting assume that the CSI reporting is triggered by DCI with format 0_1 or format 0_2 containing a non-zero CSI request field.

[0088] Time and frequency resources that may be used by the UE to report CSI are controlled and are indicated by the gNB. CSI may include channel quality indicator (CQI), precoding matrix indicator (PMI), CSI-RS resource indicator (CRI), synchronization signal (SS)/physical broadcast channel (PBCH) block resource indicator (SSBRI), layer indicator (LI), rank indicator (RI), layer-1(L1)- reference signal receiving power (RSRP), L1- signal to interference plus noise ratio (SINR) or CapabilityIndex. The corresponding definition can be found in 3GPP Technical Specification TS38.214 V18.0.0 and TS38.215 V18.0.0.

[0089] For CQI, PMI, CRI, SSBRI, LI, RI, L1-RSRP, L1-SINR, and/or CapabilityIndex reporting, a UE is configured with $N \geq 1$ Reporting Settings by higher layer parameter *CSI-ReportConfig*, $M \geq 1$ Resource Settings by higher layer parameter *CSI-ResourceConfig*, and one or two list(s) of trigger states (given by higher layer parameters *CSI-AperiodicTriggerStateList* and *CSI-SemiPersistentOnPUSCH-TriggerStateList*). Each trigger state in *CSI-AperiodicTriggerStateList* contains a list of associated *CSI-ReportConfigs* with aperiodic reporting. Each trigger state in *CSI-SemiPersistentOnPUSCH-TriggerStateList* contains one associated *CSI-ReportConfig* with semi-persistent reporting on PUSCH.

[0090] Each reporting setting *CSI-ReportConfig* is associated with a single downlink bandwidth part (BWP) (indicated by higher layer parameter *BWP-Id*) given in the

associated *CSI-ResourceConfig* for channel measurement, and it contains parameter(s) for one CSI reporting band: codebook configuration including codebook subset restriction, time-domain behavior, frequency granularity for CQI and PMI, measurement restriction configurations, and the CSI-related quantities to be reported by the UE such as LI, L1-RSRP, L1-SINR, CRI, SSBRI, and CapabilityIndex.

[0091] The time domain behavior of the *CSI-ReportConfig* is indicated by a higher layer parameter *reportConfigType* and can be set to 'aperiodic', 'semiPersistentOnPUCCH', 'semiPersistentOnPUSCH', or 'periodic'. For 'periodic' and 'semiPersistentOnPUCCH'/'semiPersistentOnPUSCH' CSI reporting, the configured periodicity and slot offset applies in the numerology of the UL BWP in which the CSI report is configured to be transmitted on. The higher layer parameter *reportQuantity* indicates the CSI-related, L1-RSRP-related, L1-SINR-related or CapabilityIndex-related quantities to report. The *reportFreqConfiguration* indicates the reporting granularity in frequency domain, including CSI reporting band and whether PMI/CQI reporting is wideband or sub-band. The *timeRestrictionForChannelMeasurements* parameter in *CSI-ReportConfig* may be configured to enable time domain restriction for channel measurements and the *timeRestrictionForInterferenceMeasurements* parameter may be configured to enable time domain restriction for interference measurements. The *CSI-ReportConfig* may also contain *CodebookConfig*, which contains configuration parameters for Type-I, Type II, Enhanced Type II CSI, or Further Enhanced Type II Port Selection including codebook subset restriction when applicable, and configurations of group-based reporting. A UE is not expected to be configured with a CSI report setting associated with a dormant DL BWP if the *reportConfigType* is set to be 'aperiodic'.

[0092] In addition, each CSI resource setting *CSI-ResourceConfig* contains a configuration of a list of $S \geq 1$ CSI resource sets (given by higher layer parameter *csi-RS-ResourceSetList*). The list includes references to either or both of non-zero-power (NZP) CSI-reference signal (RS) resource set(s) and SS/PBCH block set(s), or the list includes references to CSI-interference measurement (IM) resource set(s). Each CSI resource setting is located in the DL BWP identified by higher layer parameter *BWP-id*, and all CSI resource settings linked to a CSI report setting have the same DL BWP.

[0093] The time domain behavior of the CSI-RS resources within a CSI resource setting are indicated by higher layer parameter *resourceType* and may be set to aperiodic,

periodic, or semi-persistent. For periodic and semi-persistent CSI resource settings, when the UE is configured with *groupBasedBeamReporting-r17*, the number of CSI resource sets configured is $S=2$, otherwise the number of CSI-RS resource sets configured is limited to $S=1$. For periodic and semi-persistent CSI resource settings, the configured periodicity and slot offset is given in the numerology of its associated DL BWP, as given by *BWP-id*. When a UE is configured with multiple *CSI-ResourceConfigs* consisting the same NZP CSI-RS resource ID, the same time domain behavior may be configured for the *CSI-ResourceConfigs*. When a UE is configured with multiple *CSI-ResourceConfigs* consisting the same CSI-IM resource ID, the same time-domain behavior may be configured for the *CSI-ResourceConfigs*. All CSI resource settings linked to a CSI report setting may have the same time domain behavior.

[0094] For the reporting configurations, the UE may calculate CSI parameters (if reported) assuming the following dependencies between CSI parameters (if reported): (1) LI may be calculated conditioned on the reported CQI, PMI, RI and CRI; (2) CQI may be calculated conditioned on the reported PMI, RI and CRI; (3) PMI may be calculated conditioned on the reported RI and CRI; and (4) RI may be calculated conditioned on the reported CRI.

[0095] A reporting configuration for CSI may be aperiodic (using PUSCH), periodic (using PUCCH), or semi-persistent using PUCCH activated by MAC CE, and semi-persistent using PUSCH activated by DCI. The CSI-RS Resources may be periodic, semi-persistent, or aperiodic. In 3GPP Technical Specification TS38.214 v18.0.0, the supported combinations of CSI reporting configurations and CSI-RS resource configurations and how the CSI reporting is triggered for each CSI-RS resource configuration are shown in Table 5.2.1.4-1 specified in TS38.214 v18.0.0. Periodic CSI-RS is configured by higher layers. Semi-persistent CSI-RS is activated and deactivated as described in Clause 5.2.1.5.2 of TS38.214 v18.0.0. Aperiodic CSI-RS is configured and triggered/activated as described in Clause 5.2.1.5.1 of TS38.214 v18.0.0.

[0096] On the other side, some conclusions have been captured in TR38.843. The following lists some potential specification impacts on CSI feedback enhancement for CSI compression using two-sides model, including but not limited to fallback mode, NW/UE alignment, model input/output, UE side data collection, NW side data collection,

CSI configuration and report, feasibility and methods to support the legacy CSI reporting principles.

[0097] For example, for fallback mode, there may be potential specification impact for supporting co-existence and fallback mechanisms between AI/ML-based CSI feedback mode and legacy non-AI/ML-based CSI feedback mode.

[0098] For NW/UE alignment, there may be potential specification impact on alignment of the quantization/dequantization method and the feedback message size between network and UE, including: (1) for vector quantization scheme, the format and size of the VQ codebook, and the size and segmentation method of the CSI generation model output; and (2) for scalar quantization scheme, uniform and non-uniform quantization with format, e.g., quantization granularity, consisting of distribution of bits assigned to each float; (3) quantization alignment using 3GPP aware mechanism.

[0099] For model input/output, there may be potential specification impact on CSI output-CSI-UE and input-CSI-NW at least for precoding matrix (e.g., the precoding matrix in spatial-frequency domain or the precoding matrix represented using angular-delay domain projection). In the case of the precoding matrix represented using angular-delay domain projection, explicit channel matrix (i.e., full Tx * Rx MIMO channel) is also studied depends on the performance evaluations, where the raw channel is in spatial-frequency domain or the raw channel is in angular-delay domain.

[0100] For UE side data collection, there may be potential specification impact on the following aspects: enhancement of CSI-RS configuration to enable higher accuracy measurement, assistance information for UE data collection for categorizing the data in forms of ID for the purpose of differentiating characteristics of data due to specific configuration, scenarios, site, etc (the provision of assistance information needs to consider feasibility of disclosing proprietary information to the other side), and signaling for triggering the data collection.

[0101] For network side data collection, there may be potential specification impact on the following aspects: (1) enhancement of sounding reference signal (SRS) and/or CSI-RS measurement and/or CSI reporting to enable higher accuracy measurement; (2) contents of the ground-truth CSI, including data sample type (e.g., precoding matrix, channel matrix, etc.), data sample format including scalar quantization and/or codebook-based quantization (e.g., e-type II like), and assistance information (e.g., time stamps,

and/or cell ID, assistance information for network data collection for categorizing the data in forms of ID for the purpose of differentiating characteristics of data due to specific configuration, scenarios, site etc., and data quality indicator); (3) latency requirement for data collection; (4) signaling for triggering the data collection; (5) ground-truth CSI report for NW side data collection for model performance monitoring, including scalar quantization for ground-truth CSI, codebook-based quantization for ground-truth CSI, radio resource control (RRC) signalling and/or L1 signalling procedure to enable fast identification of AI/ML model performance, and aperiodic/semi-persistent or periodic ground-truth CSI report; (6) ground-truth CSI format *for model training*, including scalar or codebook-based quantization for ground-truth CSI. Herein, the number of layers for which the ground truth data is collected, and whether the UE or the NW determines the number of layers for ground-truth CSI data collection, are considered.

[0102] For CSI configuration and report, there may be potential specification impact on the following aspects: (1) NW configuration to determine CSI payload size, e.g., possible CSI payload size, possible rank restriction and/or other related configuration; (2) how UE determines/reports the actual CSI payload size and/or other CSI related information within constraints configured by the network; and (3) relevant uplink control information (UCI) format considering the legacy CSI reporting principle with CSI Part 1 and Part 2 as a starting point, where Part 1 has a network configured fixed size and Part 2 size is dynamic, determined by information in Part 1. For CQI determination in CSI report, if CQI in CSI report is configured, CQI is NOT calculated based on the output of CSI reconstruction part from the realistic channel estimation, instead, the CQI may be calculated based on target CSI with realistic channel measurement, or the CQI may be calculated based on target CSI with realistic channel measurement and potential adjustment, or the CQI may be calculated based on legacy codebook. Alternatively, for CQI determination in CSI report, if CQI in CSI report is configured, the CQI may be calculated based on the output of CSI reconstruction part from the realistic channel estimation. For example, the CQI may be calculated based on CSI reconstruction output if CSI reconstruction model is available at the UE and the UE can perform reconstruction model inference with potential adjustment (in this case, the CSI reconstruction part at the UE may be different comparing to the actual CSI reconstruction part used at the NW), and as another selection, the CQI may be calculated using two stage approach, and UE derive CQI using precoded CSI-RS transmitted with a reconstructed precoder.

Furthermore, feasibility of different ways to calculate CQI may be evaluated. Gap analyses between the UE side CQI calculation results and the NW side results, as well as the impact on the scheduling performance may also be evaluated. In addition, complexity of CQI calculation needs to be evaluated, including the computing complexity and potential RS/signaling overhead.

[0103] For feasibility and methods to support the legacy CSI reporting principles, there may be potential specification impact on the priority rule regarding CSI collision handling and CSI omission, the codebook subset restriction (e.g., input-CSI-NW/output-CSI-UE considered in angular-delay domain, beam restriction may be based on legacy SD basis vector-based input CSI in angular domain), and CSI processing unit.

[0104] In addition to above potential specification impacts, there are also some potential specification enhancement on: (1) CSI-RS configurations (not including CSI-RS pattern design enhancements); (2) CSI configuration (for network to indicate CSI reporting related information, e.g., gNB indication to the UE of one or more of information indicating CSI payload size, information indicating quantization method/granularity, rank restriction, and other payload related aspects); (3) CSI reporting configurations (for UE determination/reporting of the actual CSI payload size, UE reports related information as configured by the NW); (4) CSI report UCI mapping/priority/omission; and (5) CSI processing procedures.

[0105] For better support the CSI compression based CSI reporting with two-sided AI/ML models, the legacy CSI reporting framework need to be further enhanced.

[0106] However, even above legacy CSI reporting framework could be enhanced to support the CSI compression based CSI reporting with two-sided AI/ML models, corresponding specification effort may be larger since lots of signaling may need to be introduced or be enhanced and UE behavior may be complicated.

[0107] On the other hand, AI/ML based CSI prediction/beam prediction and the like may also be specified in future NR releases and may be further supported in 6th generation (6G). The required CSI resource settings, reporting contents and additional AI/ML related configurations may be different from those in Release 17.

[0108] A new CSI reporting framework for the UE and NW with AI/ML capability is needed to better support AI/ML based functionalities (which may also be referred to as AI/ML functionalities below) in UE and/or NW.

[0109] The present disclosure proposes a solution of CSI reporting framework to support CSI reporting with AI/ML functionalities, for example, an AI/ML functionality for CSI compression, an AI/ML functionality for CSI prediction, an AI/ML functionality for spatial domain beam prediction, or an AI/ML functionality for temporal domain beam prediction. In this solution, the UE may report its capability on support of one or more AI/ML functionalities to the NW, e.g., a base station. Based on the received UE capacity, the NW may configure AI/ML CSI report configuration(s) for the UE and send the configuration signaling to the UE. The UE may perform AI/ML related operations based on the received AI/ML CSI report configuration(s). In this way, AI/ML functionalities may be better supported at both UE side and the NW side. Furthermore, the CSI reporting framework proposed in present disclosure is more appropriate to support CSI reporting with AI/ML ability and is more friendly for the AI/ML operation towards 6G.

[0110] Aspects of the present disclosure are described in the context of a wireless communications system.

[0111] FIG. 1 illustrates an example of a wireless communications system 100 that supports CSI reporting in accordance with aspects of the present disclosure. The wireless communications system 100 may include one or more network entities 102 (also referred to as network equipment (NE)), one or more UEs 104, a core network 106, and a packet data network 108. The wireless communications system 100 may support various radio access technologies. In some implementations, the wireless communications system 100 may be a 4G network, such as an LTE network or an LTE-advanced (LTE-A) network. In some other implementations, the wireless communications system 100 may be a 5G network, such as an NR network. In other implementations, the wireless communications system 100 may be a combination of a 4G network and a 5G network, or other suitable radio access technology including institute of electrical and electronics engineers (IEEE) 802.11 (Wi-Fi), IEEE 802.16 (WiMAX), IEEE 802.20. The wireless communications system 100 may support radio access technologies beyond 5G. Additionally, the wireless communications system 100 may support technologies, such as time division multiple

access (TDMA), frequency division multiple access (FDMA), or code division multiple access (CDMA), etc.

[0112] The one or more network entities 102 may be dispersed throughout a geographic region to form the wireless communications system 100. One or more of the network entities 102 described herein may be or include or may be referred to as a network node, a base station, a network element, a radio access network (RAN), a base transceiver station, an access point, a NodeB, an eNodeB (eNB), a next-generation NodeB (gNB), or other suitable terminology. A network entity 102 and a UE 104 may communicate via a communication link 110, which may be a wireless or wired connection. For example, a network entity 102 and a UE 104 may perform wireless communication (e.g., receive signaling, transmit signaling) over a Uu interface.

[0113] A network entity 102 may provide a geographic coverage area 112 for which the network entity 102 may support services (e.g., voice, video, packet data, messaging, broadcast, etc.) for one or more UEs 104 within the geographic coverage area 112. For example, a network entity 102 and a UE 104 may support wireless communication of signals related to services (e.g., voice, video, packet data, messaging, broadcast, etc.) according to one or multiple radio access technologies. In some implementations, a network entity 102 may be moveable, for example, a satellite associated with a non-terrestrial network. In some implementations, different geographic coverage areas 112 associated with the same or different radio access technologies may overlap, but the different geographic coverage areas 112 may be associated with different network entities 102. Information and signals described herein may be represented using any of a variety of different technologies and techniques. For example, data, instructions, commands, information, signals, bits, symbols, and chips that may be referenced throughout the description may be represented by voltages, currents, electromagnetic waves, magnetic fields or particles, optical fields or particles, or any combination thereof.

[0114] The one or more UEs 104 may be dispersed throughout a geographic region of the wireless communications system 100. A UE 104 may include or may be referred to as a mobile device, a wireless device, a remote device, a remote unit, a handheld device, or a subscriber device, or some other suitable terminology. In some implementations, the UE 104 may be referred to as a unit, a station, a terminal, or a client, among other examples. Additionally, or alternatively, the UE 104 may be referred to as an internet-of-

things (IoT) device, an internet-of-everything (IoE) device, or machine-type communication (MTC) device, among other examples. In some implementations, a UE 104 may be stationary in the wireless communications system 100. In some other implementations, a UE 104 may be mobile in the wireless communications system 100.

[0115] The one or more UEs 104 may be devices in different forms or having different capabilities. Some examples of UEs 104 are illustrated in FIG. 1. A UE 104 may be capable of communicating with various types of devices, such as the network entities 102, other UEs 104, or network equipment (e.g., the core network 106, the packet data network 108, a relay device, an integrated access and backhaul (IAB) node, or another network equipment), as shown in FIG. 1. Additionally, or alternatively, a UE 104 may support communication with other network entities 102 or UEs 104, which may act as relays in the wireless communications system 100.

[0116] A UE 104 may also be able to support wireless communication directly with other UEs 104 over a communication link 114. For example, a UE 104 may support wireless communication directly with another UE 104 over a device-to-device (D2D) communication link. In some implementations, such as vehicle-to-vehicle (V2V) deployments, vehicle-to-everything (V2X) deployments, or cellular-V2X deployments, the communication link 114 may be referred to as a sidelink. For example, a UE 104 may support wireless communication directly with another UE 104 over a PC5 interface.

[0117] A network entity 102 may support communications with the core network 106, or with another network entity 102, or both. For example, a network entity 102 may interface with the core network 106 through one or more backhaul links 116 (e.g., via an S1, N2, N2, or another network interface). The network entities 102 may communicate with each other over the backhaul links 116 (e.g., via an X2, Xn, or another network interface). In some implementations, the network entities 102 may communicate with each other directly (e.g., between the network entities 102). In some other implementations, the network entities 102 may communicate with each other or indirectly (e.g., via the core network 106). In some implementations, one or more network entities 102 may include subcomponents, such as an access network entity, which may be an example of an access node controller (ANC). An ANC may communicate with the one or more UEs 104 through one or more other access network transmission entities, which

may be referred to as a radio heads, smart radio heads, or transmission-reception points (TRPs).

[0118] In some implementations, a network entity 102 may be configured in a disaggregated architecture, which may be configured to utilize a protocol stack physically or logically distributed among two or more network entities 102, such as an integrated access backhaul (IAB) network, an open radio access network (O-RAN) (e.g., a network configuration sponsored by the O-RAN Alliance), or a virtualized RAN (vRAN) (e.g., a cloud RAN (C-RAN)). For example, a network entity 102 may include one or more of a CU, a DU, a radio unit (RU), a RAN intelligent controller (RIC) (e.g., a near-real time RIC (Near-RT RIC), a non-real time RIC (Non-RT RIC)), a service management and orchestration (SMO) system, or any combination thereof.

[0119] An RU may also be referred to as a radio head, a smart radio head, a remote radio head (RRH), a remote radio unit (RRU), or a transmission reception point (TRP). One or more components of the network entities 102 in a disaggregated RAN architecture may be co-located, or one or more components of the network entities 102 may be located in distributed locations (e.g., separate physical locations). In some implementations, one or more network entities 102 of a disaggregated RAN architecture may be implemented as virtual units (e.g., a virtual CU (VCU), a virtual DU (VDU), a virtual RU (VRU)).

[0120] Split of functionality between a CU, a DU, and an RU may be flexible and may support different functionalities depending upon which functions (e.g., network layer functions, protocol layer functions, baseband functions, radio frequency functions, and any combinations thereof) are performed at a CU, a DU, or an RU. For example, a functional split of a protocol stack may be employed between a CU and a DU such that the CU may support one or more layers of the protocol stack and the DU may support one or more different layers of the protocol stack. In some implementations, the CU may host upper protocol layer (e.g., a layer 3 (L3), a layer 2 (L2)) functionality and signaling (e.g., radio resource control (RRC), service data adaptation protocol (SDAP), packet data convergence protocol (PDCP)). The CU may be connected to one or more DUs or RUs, and the one or more DUs or RUs may host lower protocol layers, such as a layer 1 (L1) (e.g., physical (PHY) layer) or an L2 (e.g., radio link control (RLC) layer, medium access control (MAC) layer) functionality and signaling, and may each be at least partially controlled by the CU 160.

[0121] Additionally, or alternatively, a functional split of the protocol stack may be employed between a DU and an RU such that the DU may support one or more layers of the protocol stack and the RU may support one or more different layers of the protocol stack. The DU may support one or multiple different cells (e.g., via one or more RUs). In some implementations, a functional split between a CU and a DU, or between a DU and an RU may be within a protocol layer (e.g., some functions for a protocol layer may be performed by one of a CU, a DU, or an RU, while other functions of the protocol layer are performed by a different one of the CU, the DU, or the RU).

[0122] A CU may be functionally split further into CU control plane (CU-CP) and CU user plane (CU-UP) functions. A CU may be connected to one or more DUs via a midhaul communication link (e.g., F1, F1-c, F1-u), and a DU may be connected to one or more RUs via a fronthaul communication link (e.g., open fronthaul (FH) interface). In some implementations, a midhaul communication link or a fronthaul communication link may be implemented in accordance with an interface (e.g., a channel) between layers of a protocol stack supported by respective network entities 102 that are in communication via such communication links.

[0123] The core network 106 may support user authentication, access authorization, tracking, connectivity, and other access, routing, or mobility functions. The core network 106 may be an evolved packet core (EPC), or a 5G core (5GC), which may include a control plane entity that manages access and mobility (e.g., a mobility management entity (MME), an access and mobility management functions (AMF)) and a user plane entity that routes packets or interconnects to external networks (e.g., a serving gateway (S-GW), a packet data network (PDN) gateway (P-GW), or a user plane function (UPF)). In some implementations, the control plane entity may manage non-access stratum (NAS) functions, such as mobility, authentication, and bearer management (e.g., data bearers, signal bearers, etc.) for the one or more UEs 104 served by the one or more network entities 102 associated with the core network 106.

[0124] The core network 106 may communicate with the packet data network 108 over one or more backhaul links 116 (e.g., via an S1, N2, N2, or another network interface). The packet data network 108 may include an application server 118. In some implementations, one or more UEs 104 may communicate with the application server 118. A UE 104 may establish a session (e.g., a protocol data unit (PDU) session, or the like)

with the core network 106 via a network entity 102. The core network 106 may route traffic (e.g., control information, data, and the like) between the UE 104 and the application server 118 using the established session (e.g., the established PDU session). The PDU session may be an example of a logical connection between the UE 104 and the core network 106 (e.g., one or more network functions of the core network 106).

[0125] In the wireless communications system 100, the network entities 102 and the UEs 104 may use resources of the wireless communications system 100 (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers)) to perform various operations (e.g., wireless communications). In some implementations, the network entities 102 and the UEs 104 may support different resource structures. For example, the network entities 102 and the UEs 104 may support different frame structures. In some implementations, such as in 4G, the network entities 102 and the UEs 104 may support a single frame structure. In some other implementations, such as in 5G and among other suitable radio access technologies, the network entities 102 and the UEs 104 may support various frame structures (i.e., multiple frame structures). The network entities 102 and the UEs 104 may support various frame structures based on one or more numerologies.

[0126] One or more numerologies may be supported in the wireless communications system 100, and a numerology may include a subcarrier spacing and a cyclic prefix. A first numerology (e.g., $\mu=0$) may be associated with a first subcarrier spacing (e.g., 15 kHz) and a normal cyclic prefix. In some implementations, the first numerology (e.g., $\mu=0$) associated with the first subcarrier spacing (e.g., 15 kHz) may utilize one slot per subframe. A second numerology (e.g., $\mu=1$) may be associated with a second subcarrier spacing (e.g., 30 kHz) and a normal cyclic prefix. A third numerology (e.g., $\mu=2$) may be associated with a third subcarrier spacing (e.g., 60 kHz) and a normal cyclic prefix or an extended cyclic prefix. A fourth numerology (e.g., $\mu=3$) may be associated with a fourth subcarrier spacing (e.g., 120 kHz) and a normal cyclic prefix. A fifth numerology (e.g., $\mu=4$) may be associated with a fifth subcarrier spacing (e.g., 240 kHz) and a normal cyclic prefix.

[0127] A time interval of a resource (e.g., a communication resource) may be organized according to frames (also referred to as radio frames). Each frame may have a duration, for example, a 10 millisecond (ms) duration. In some implementations, each

frame may include multiple subframes. For example, each frame may include 10 subframes, and each subframe may have a duration, for example, a 1 ms duration. In some implementations, each frame may have the same duration. In some implementations, each subframe of a frame may have the same duration.

[0128] Additionally or alternatively, a time interval of a resource (e.g., a communication resource) may be organized according to slots. For example, a subframe may include a number (e.g., quantity) of slots. The number of slots in each subframe may also depend on the one or more numerologies supported in the wireless communications system 100. For instance, the first, second, third, fourth, and fifth numerologies (i.e., $\mu=0$, $\mu=1$, $\mu=2$, $\mu=3$, $\mu=4$) associated with respective subcarrier spacings of 15 kHz, 30 kHz, 60 kHz, 120 kHz, and 240 kHz may utilize a single slot per subframe, two slots per subframe, four slots per subframe, eight slots per subframe, and 16 slots per subframe, respectively. Each slot may include a number (e.g., quantity) of symbols (e.g., OFDM symbols). In some implementations, the number (e.g., quantity) of slots for a subframe may depend on a numerology. For a normal cyclic prefix, a slot may include 14 symbols. For an extended cyclic prefix (e.g., applicable for 60 kHz subcarrier spacing), a slot may include 12 symbols. The relationship between the number of symbols per slot, the number of slots per subframe, and the number of slots per frame for a normal cyclic prefix and an extended cyclic prefix may depend on a numerology. It should be understood that reference to a first numerology (e.g., $\mu=0$) associated with a first subcarrier spacing (e.g., 15 kHz) may be used interchangeably between subframes and slots.

[0129] In the wireless communications system 100, an electromagnetic (EM) spectrum may be split, based on frequency or wavelength, into various classes, frequency bands, frequency channels, etc. By way of example, the wireless communications system 100 may support one or multiple operating frequency bands, such as frequency range designations FR1 (410 MHz – 7.125 GHz), FR2 (24.25 GHz – 52.6 GHz), FR3 (7.125 GHz – 24.25 GHz), FR4 (52.6 GHz – 114.25 GHz), FR4a or FR4-1 (52.6 GHz – 71 GHz), and FR5 (114.25 GHz – 300 GHz). In some implementations, the network entities 102 and the UEs 104 may perform wireless communications over one or more of the operating frequency bands. In some implementations, FR1 may be used by the network entities 102 and the UEs 104, among other equipment or devices for cellular communications traffic (e.g., control information, data). In some implementations, FR2 may be used by the

network entities 102 and the UEs 104, among other equipment or devices for short-range, high data rate capabilities.

[0130] FR1 may be associated with one or multiple numerologies (e.g., at least three numerologies). For example, FR1 may be associated with a first numerology (e.g., $\mu=0$), which includes 15 kHz subcarrier spacing; a second numerology (e.g., $\mu=1$), which includes 30 kHz subcarrier spacing; and a third numerology (e.g., $\mu=2$), which includes 60 kHz subcarrier spacing. FR2 may be associated with one or multiple numerologies (e.g., at least 2 numerologies). For example, FR2 may be associated with a third numerology (e.g., $\mu=2$), which includes 60 kHz subcarrier spacing; and a fourth numerology (e.g., $\mu=3$), which includes 120 kHz subcarrier spacing.

[0131] FIG. 2 illustrates an example of signalling procedure 200 for CSI reporting in accordance with aspects of the present disclosure. The signalling procedure 200 involves a base station 102 and a UE 104, the base station 102 may also be referred to as NW 102, gNB 102 or the like.

[0132] As shown in FIG. 2, at step 202, the UE 104 may transmit capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities to the base station 102. The capability information may be determined by the UE 104 per UE, per band, or per band combination. That is, the UE 104 may report its capability information for the all bands to the base station 102, or the UE 104 may report its capability information on each band to the base station 102. Alternatively, the UE 104 may report its capability information on each band combination to the base station 102. Furthermore, an AI/ML functionality may be associated with one or more AI/ML models, and the UE 104 may transmit its capability information per functionality or per model, that is, for example, the UE 104 may transmit multiple capability information, and each capability information corresponds to one AI/ML functionality or one AI/ML model.

[0133] In some embodiments of present disclosure, the one or more AI/ML functionalities of the UE 104 may include at least one of: an AI/ML functionality for CSI compression, an AI/ML functionality for CSI prediction, an AI/ML functionality for spatial domain beam prediction, and an AI/ML functionality for temporal domain beam prediction. In the following description, the AI/ML functionality for CSI prediction and the AI/ML functionality for CSI compression will be used as examples for explaining embodiments of present disclosure. However, similar principles are applicable to other

AI/ML functionalities, such as the AI/ML functionality for spatial domain beam prediction and the AI/ML functionality for temporal domain beam prediction, but detailed description thereof is omitted for concision.

[0134] In some embodiments of present disclosure, as an example, the capability information may include at least one of a first capability information on the AI/ML functionality for CSI compression and a second capability information on the AI/ML functionality for CSI prediction. That is, the UE 104 may report or indicate, to the base station 102, its capabilities in corresponding feature groups on the support of CSI compression using two-sided model and the support of CSI prediction using a UE side model.

[0135] In some embodiments of present disclosure, the first capability information may include a first information on CSI compression reporting type (including at least one of per-rank CSI reporting and per-layer CSI reporting), a second information on a number of uplink control information (UCI) bits for a compressed CSI report, and a third information on whether a vector-based quantization scheme or a scalar based quantization scheme for compressed CSI feedback is supported by the UE 104. For example, for a feature group of supporting CSI compression using two-sided model, the UE 104 may indicate, in the first capability information, CSI compression reporting type(s) it supports, which may be per rank reporting or per layer reporting. For per rank reporting, the UE 104 may further indicate, in the first capability information, one or multiple candidate numbers of required UCI bits for a compressed CSI report corresponding to each of the supported rank. Similarly, for per layer reporting, the UE 104 may further indicate, in the first capability information, a candidate number of required UCI bits for a compressed CSI report corresponding to each transmission layer. In addition, for each CSI compression reporting type, the UE 104 may further indicate, in the first capability information, whether it supports a vector-based quantization scheme or a scalar based quantization scheme for compressed CSI feedback.

[0136] Moreover, the second capability information may include a fourth information on CSI type(s) to be predicted, and a fifth information on a relation between a number of historical CSI-reference signal (RS) receptions within a measurement time window (which may be used to obtain a number of input CSIs for an AI/ML model at UE 104) and a number of predicted CSIs in a prediction time window. For example, the relation

may be that 5 CSI-RS receptions in 5ms as inputs for the AI/ML model to predict 2 CSIs in future 5ms. Herein, the CSI type(s) to be predicted may include at least one of rank indicator (RI), precoding matrix indicator (PMI), layer indicator (LI), and channel quality indicator (CQI). For example, for a feature group of supporting CSI prediction using a UE side model, the UE 104 may indicate, in the second capability information, CSI type(s) it can be predicted, e.g., RI, PMI, LI and CQI. In the case that the CSI type(s) to be predicted includes at least one of the PMI and the CQI, the second capability information may further include a sixth information on whether at least one of wideband CSI prediction and subband CSI prediction is supported by the UE 104, and a seventh information on a codebook type for PMI supported by the UE 104. That is, if the UE 104 can perform PMI and/or CQI prediction, the UE 104 may further indicate, in the second capability information, the support of wideband and/or sub-band CSI prediction, and the supported codebook type for PMI report, e.g., Type-I single-panel codebook, Type-II codebook or Type-II port selection codebook, which are specified in 3GPP TS38.214 V18.0.0. In addition, the UE may further indicate, in the second capability information, the relation between the number of historical CSI-RS receptions within a measurement window, which are used to obtain the number of input CSIs for the model, and the number of predicted CSIs in a prediction window. For example, 5 CSI-RS receptions in 5ms as the model input to predict 2 CSIs in future 5ms.

[0137] Referring to FIG. 2 again, the base station 102 may receive the capability information transmitted by the UE 104 at step 202, and transmit, at step 204, at least one CSI report configuration to the UE 104. The at least one CSI report configuration may at least include at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model. Herein, each AI/ML CSI report configuration may be associated with an AI/ML functionality or an AI/ML model, and an AI/ML functionality may be associated with one or more AI/ML models. Furthermore, the at least one CSI report configuration may further include other type(s) of CSI report configuration, such as legacy CSI report configuration(s), layer 1 and/or layer 2 (L1/L2)-triggered mobility (LTM) CSI report configuration(s)(i.e., CSI report configuration(s) supporting LTM), and the like, which will be described below later.

[0138] The UE 104 may receive the at least one CSI report configuration transmitted by the base station 102 at step 204, and may perform some AI/ML related functionalities or AI/ML related operations based on the at least one CSI report configuration, such as

the AI/ML functionality for CSI compression, the AI/ML functionality for CSI prediction, the AI/ML functionality for spatial domain beam prediction, the AI/ML functionality for temporal domain beam prediction, and/or the like. Some exemplary description regarding the AI/ML CSI report configuration and the related UE behavior are described in details below.

[0139] In some embodiments of present disclosure, each of the at least one AI/ML CSI report configuration may be transmitted via a preset higher layer information element (IE) configured by the base station 102, for example, a higher layer IE *AIML-CSI-ReportConfig*.

[0140] For example, the base station 102 may configure a list of AI/ML-based CSI report configurations (i.e., CSI report configurations with AI/ML ability, which may also be expressed as AI/ML CSI report configurations), e.g., by a higher layer parameter *AIML-CSI-ReportConfigToAddModList*. Each of the AI/ML CSI report configurations (e.g., configured by a higher layer information element *AIML-CSI-ReportConfig* contained in the *AIML-CSI-ReportConfigToAddModList*) may have an ID, e.g., *AIML-CSI-ReportConfigId*, and may be further associated with an AI/ML functionality ID or an AI/ML model ID for the UE 104 to determine an appropriate UE-side AI/ML model to obtain a required CSI report. Furthermore, an AI/ML CSI report configuration may be associated with a legacy CSI report configuration (which may also referred to as non-AI/ML CSI report configuration), which may be associated with the same CSI resource setting for channel measurement as the AI/ML CSI report configuration, for fallback operation without AI/ML operations.

[0141] In some embodiments of present disclosure, an AI/ML CSI report configuration may include at least one of: a configuration identifier (ID) indicating the AI/ML CSI report configuration identity, a resource information on CSI resource for channel measurement and/or CSI resource for interference measurement, an AI/ML functionality or model information indicating a AI/ML functionality or model associated with the AI/ML CSI report configuration, a usage information indicating the usage of the AI/ML CSI report configuration, a reporting type for the AI/ML CSI report configuration, a rank subset restriction indicating an allowed rank indicator (RI) report for the AI/ML CSI report configuration, and an associated CSI report configuration for fallback operation. The associated CSI report configuration of the AI/ML CSI report configuration

may indicate a non-AI/ML CSI report configuration for fallback operation without AI/ML operation, and the reporting type for the AI/ML CSI report configuration may include one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic.

[0142] As an example, each *AIML-CSI-ReportConfig* may at least include the following parameters: *AIML-CSI-ReportConfigId* (corresponding to the configuration ID described above), resources for channel measurement and/or resources for interference measurement (corresponding to the resource information described above), associated AI/ML functionality (corresponding to the AI/ML functionality or model information described above), usage and the report content (corresponding to the usage information described above), report type (corresponding to the reporting type described above), rank subset restriction (corresponding to the rank subset restriction described above), and associated *CSI-ReportConfig* for fallback operation (corresponding to the associated CSI report configuration described above).

[0143] The parameter *AIML-CSI-ReportConfigId* may be used to determine an AI/ML CSI Report configuration identity, in other words, identify an AI/ML CSI Report configuration. The resources for channel measurement may be used to indicate a resource setting containing one or more CSI-RS resource set(s) or SSB resource set(s) for channel measurement, and the resources for interference measurement may be used to indicate a resource setting containing one or more ZP CSI-RS resource set(s) or NZP CSI-RS resource set(s) for interference measurement.

[0144] Furthermore, in some embodiments of present disclosure, for the AI/ML functionality or model information, in case that there are more than one AI/ML functionality identified at the UE 104, the AI/ML functionality or model information may include a functionality ID of an AI/ML functionality associated with the corresponding AI/ML CSI report configuration, while in case that there is only one AI/ML functionality identified at the UE 104, the AI/ML functionality or model information may include an AI/ML model ID for an AI/ML functionality associated with the corresponding AI/ML CSI report configuration, or it may indicate to enable the AI/ML functionality associated with the AI/ML CSI report configuration. For example, each of the *AIML-CSI-ReportConfig* may be associated with a functionality ID when more than one AI/ML functionality are identified at the UE side. The AI/ML model used for inference may be

determined by the UE 104 itself. If only one AI/ML functionality with more than one AI/ML models is identified at the UE side, the base station 102 may directly associate an AI/ML model ID of an AI/ML model among the AI/ML models to the *AIML-CSI-ReportConfig* or enable a certain AI/ML functionality, such as an AI/ML functionality for CSI compression, an AI/ML functionality for CSI prediction, an AI/ML functionality for spatial domain beam prediction, or an AI/ML functionality for temporal domain beam prediction.

[0145] In some embodiments of present disclosure, the usage of an AI/ML CSI report configuration may include at least one of data collection for model training, model inference, and performance monitoring. The model inference may include at least one of model inference for CSI compression, model inference for CSI prediction, model inference for spatial domain beam prediction, and model inference for temporal domain beam prediction. Depends on different usage of an AI/ML CSI report configuration, an AI/ML CSI report configured based on the AI/ML CSI report configuration and to be transmitted by the UE 104 to the base station 102 may include different report content, which will be described in detail below.

[0146] As a first example, in some embodiments of present disclosure, in the case that an AI/ML CSI report configuration is configured for data collection for model training for two-sided model for CSI compression, a first CSI obtained by using an AI/ML model at the UE together with a second CSI obtained without using any AI/ML model at the UE, which is used as a model input for the AI/ML model at the UE 104, may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration. For example, the base station 102 may configure the UE 104 to report both the compressed CSI (i.e., CSI generated with AI/ML operation) and the original CSI (i.e., legacy CSI which is generated without AI/ML operation), e.g., Rel-15 Type II CSI or Rel-16 eType II CSI, in the same AI/ML CSI report. One example for corresponding UCI report format is provided in Table 1 below, where only one CSI-RS resource Indicator (CRI), RI, layer indicator (LI) and/or channel quality indicator (CQI) are reported in the AI/ML CSI report, while two PMI parts are reported, including AI PMI (i.e., the compressed CSI) part and non-AI PMI (i.e., the original CSI, e.g., Rel-16 eType II PMI) part. The AI PMI and the non-AI PMI may be calculated based on reported CRI and RI. The LI and CQI may be calculated based on reported non-AI PMI.

Table 1 Mapping order of CSI fields of one CSI report containing both AI PMI and non-AI PMI for data collection (assuming all are reported per wideband)

CSI report number	CSI field	Bitwidth
CSI report# n	CSI-RS resource indicator (CRI)	$\lceil \log_2(K_s^{CSI-RS}) \rceil$
	Rank indicator (RI)	$\lceil \log_2(n_{RI}) \rceil$
	Layer indicator (LI)	$\lceil \log_2(v) \rceil$
	AI PMI part: Compressed CSI corresponding to the indicated RI	$N_{UCI-AI-PMI}$
	Non-AI PMI part: Rel-16 eType-II PMI	As that in Rel-16

[0147] In Table 1, K_s^{CSI-RS} is the number of CSI-RS resources within the CSI-RS resource set associated with the *AI/ML-CSI-ReportConfig*, n_{RI} represents the max value of reported RI, and v represents the number of reported layers. $N_{UCI-AI-PMI}$ represents the number of UCI bits for the AI PMI.

[0148] Furthermore, as a second example, in the case that an AI/ML CSI report configuration is configured for inference for CSI compression, and a number of uplink control information (UCI) bits for a compressed CSI report is determined by the UE 104, the number of UCI bits for the compressed CSI report may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and a compressed CSI may be included in a second part of the AI/ML CSI report. For example, A higher layer parameter *reportQuantity* may be configured to indicate CSI report contents for this AI/ML CSI report, such as CRI-RI-PMI (or, eigenvectors), CRI-RI-PMI (or, eigenvectors)-CQI or CRI-RI-LI-PMI (or, eigenvectors)-CQI. However, different from legacy codebook-based PMI feedback, no common codebook will be specified for AI/ML based CSI compression, the compressed CSI may only be explained by a corresponding CSI generation model and a corresponding CSI reconstruction model. The corresponding UCI bits may be configured by the base station 102 according to the reported UE capability via the capability information, or may be directly indicated by the UE 104. As an example, the base station 102 may indicate a CSI compression reporting type of a compressed CSI report, e.g., which may be set as per-rank CSI reporting or per-layer CSI reporting, as mentioned above. For the per rank reporting, the base station 102 may further indicate the number of UCI bits for the compressed CSI report corresponding to each supported rank, and for the per layer reporting, the base station 102 may further indicate the number of UCI bits for the compressed CSI report corresponding to each

layer (e.g., each transmission layer). As another example, the number of UCI bits for a compressed CSI report (e.g., AI PMI report) may be determined by the UE 104 and reported in the Part-1 CSI report (e.g., a first part of an AI/ML CSI report), and the compressed CSI, i.e., the AI PMI, may be reported in Part-2 CSI report (e.g., a second part of the same AI/ML CSI report). As a result, CRI, RI, LI and the number of UCI bits for AI PMI may be reported in Part-1 CSI report, while the AI PMI may be reported in Part-2 CSI report.

[0149] In addition, as a third example, in the case that an AI/ML CSI report configuration is configured for inference for CSI prediction, a number of time instance of a predicted CSI and a predicted CSI corresponding to a first time instance may be included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and other predicted CSI corresponding to other time instance may be included in a second part of the AI/ML CSI report. For example, for time domain CSI/beam prediction, a higher layer parameter *reportQuantity* may be configured to indicate report content for a corresponding AI/ML CSI report, e. g., predicted CSI report contents and the number of time instance for prediction, according to the UE capability reported in the capability information. Alternatively, the number of time instance of the predicted CSI may be determined by the UE 104 and indicated in the Part-1 CSI report (e.g., a first part of an AI/ML CSI report) along with a first predicted CSI corresponding to the first time instance, and the predicted CSIs other than the first predicted CSI may be indicated in Part-2 CSI (e.g., a second part of the same AI/ML CSI report).

[0150] As a fourth example, in the case that an AI/ML CSI report configuration is configured for performance monitoring, a performance gap between a reported CSI (i.e. AI/ML-based CSI or AI/ML CSI) and the actual CSI of a same CSI-RS resource on a same transmission occasion may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration. The base station 102 may indicate the UE 104 to report the performance gap, e.g., a MSE level between the reported CSI and the actual CSI, where the reported CSI may be eigenvector(s) of corresponding CSI reconstruction model output and the actual CSI may be eigenvector(s) of an estimated channel matrix for CSI compression obtained by a CSI-RS resource in a same transmission occasion. Alternatively, in this case, a ground-truth CSI corresponding to a CSI-RS resource or a CSI-RS resource set with a certain quantization scheme may be included in an AI/ML CSI report configured based on the AI/ML CSI report configuration, e.g., the base station

102 may indicate the UE 104 to report the ground-truth CSI corresponding to a certain CSI-RS resource or CSI-RS resource set with a specific quantization scheme.

[0151] In above examples, the base station 102 may indicate the report content to the UE 104 via an IE, or the report content may be determined based on corresponding usage of the AI/ML CSI report configuration or based on other ways.

[0152] In some embodiments of present disclosure, the report type may be set as periodic, semiPersistentOnPUCCH, semiPersistentOnPUSCH and aperiodic, and for semiPersistentOnPUCCH, PUCCH resources may be provided as well. Moreover, the rank subset restriction may be a bitmap to indicate allowed RI report for corresponding *AIML-CSI-ReportConfig*.

[0153] The associated *CSI-ReportConfig* for fallback operation may be an identifier of a legacy CSI report configuration used for performance monitoring and/or fallback. In some embodiments of present disclosure, in the case that the UE 104 determines to perform a fallback operation without AI/ML operation based on an AI/ML CSI report configuration, the UE 104 may drop an AI/ML CSI report corresponding to the AI/ML CSI report configuration, and transmit a non-AI/ML CSI report corresponding to an associated non-AI/ML CSI report configuration for the AI/ML CSI report configuration. The same CSI resources for channel and/or interference measurement for the AI/ML CSI report configuration may be used for the associated non-AI/ML CSI report configuration. That is, for fallback operation, the UE 104 may expect that the same CSI resource sets are configured for the associated *CSI-ReportConfig* and the *AIML-CSI-ReportConfig*. With this association, when the UE 104 determines to perform fallback operation, the UE 104 may drop a CSI corresponding to the *AIML-CSI-ReportConfig* and transmit a CSI corresponding to the associated *CSI-ReportConfig* for the *AIML-CSI-ReportConfig*.

[0154] In some embodiments of present disclosure, the at least one CSI report configuration transmitted from the base station 102 to the UE 104 may further include at least one non-AI/ML CSI report configuration, for example, at least one legacy CSI report configuration or other type of CSI report configuration.

[0155] In some embodiments of present disclosure, the at least one CSI report configuration transmitted from the base station 102 to the UE 104 may include at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink control channel (PUCCH), and/or at least one non-AI/ML CSI report configuration

for SP CSI report on PUCCH. In some embodiments of present disclosure, for activation/deactivation of a SP AI/ML CSI report on PUCCH, the UE 104 may receive a media access control (MAC) control element (CE) from the base station 102.

[0156] FIGS. 3A to 3C illustrate examples of SP AI/ML CSI report on PUCCH activation/deactivation MAC CE in accordance with aspects of the present disclosure. For each of MAC CE format as shown in FIGS. 3A to 3C, Serving Cell ID field may indicate an identity of a serving cell for which the MAC CE applies, and BWP ID field may indicate a uplink (UL) bandwidth part (BWP) for which the MAC CE applies.

[0157] In some embodiments of present disclosure, the MAC CE may include an AI/ML field indicating whether the MAC CE is used for activating at least one AI/ML CSI report configuration for SP CSI report on PUCCH or at least one non-AI/ML CSI report configuration for SP CSI report on PUCCH.

[0158] As shown in FIG. 3A, an AI/ML field may be introduced by reusing some reserved bits to indicate the activated SP AI/ML CSI report on PUCCH. The AI/ML field may include a plurality of bits (e.g., T_0 to T_3 bits), each of which indicates whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated. In other words, the legacy (e.g., non-AI/ML) SP CSI report on PUCCH activation/deactivation MAC CE may be enhanced by adding a bitmap to indicate the activation/deactivation status for each AI/ML CSI report configuration for SP CSI report on PUCCH.

[0159] For the MAC CE format in FIG. 3A, the logical channel identify (LCID) used for the SP CSI report on PUCCH activation/deactivation MAC CE may be reused to identify this enhanced MAC CE. The bitmap S_0, S_1, S_2, S_3 may indicate activation/deactivation status of SP CSI report configurations (e.g. non-AI/ML CSI report configurations for SP CSI report on PUCCH) within *csi-ReportConfigToAddModList*, S_0 refers to the CSI report configuration which includes PUCCH resources for SP CSI reporting in the indicated BWP and has the lowest *CSI-ReportConfigId* within the list with type set to *semiPersistentOnPUCCH*, S_1 refers to the CSI report configuration which includes PUCCH resources for SP CSI reporting in the indicated BWP and has the second lowest *CSI-ReportConfigId*, and so on. The S_i field is set to 1 to indicate that the corresponding SP CSI report configuration shall be activated. The S_i field is set to 0 to indicate that the corresponding SP CSI report configuration shall be deactivated. The

bitmap T_0, T_1, T_2, T_3 may indicate activation/deactivation status of SP AI/ML CSI report configurations (e.g., AI/ML CSI report configurations for SP CSI report on PUCCH) within *AIML-CSI-ReportConfigToAddModList*, T_0 refers to the CSI report configuration which includes PUCCH resources for SP AI/ML CSI reporting in the indicated BWP and has the lowest *AI/ML-CSI-ReportConfigId* within the list with type set to *semiPersistentOnPUCCH*, T_1 refers to the report configuration which includes PUCCH resources for SP AI/ML CSI reporting in the indicated BWP and has the second lowest *AI/ML-CSI-ReportConfigId*, and so on. The T_i field is set to 1 to indicate that the corresponding SP AI/ML CSI report configuration shall be activated. The T_i field is set to 0 to indicate that the corresponding SP AI/ML CSI report configuration shall be deactivated.

[0160] Alternatively, as shown in FIG. 3B, the AI/ML field may include one bit (i.e. the AI/ML bits in FIG. 3B) indicating whether a predetermined field (e.g., S_0 to S_3 bits) in the MAC CE is used for AI/ML or non-AI/ML. In the case that the one bit has a first value, each bit in the predetermined field may indicate whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated, but in the case that the one bit has a second value different from the first value, each bit in the predetermined field may indicate whether a corresponding non-AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated. In other words, the legacy (e.g., non-AI/ML) SP CSI report on PUCCH activation/deactivation MAC CE may be enhanced by adding a field (i.e., the AI/ML field in FIG. 3B, one bit) to indicate whether the MAC CE is used to activate/deactivate non-AI/ML CSI report configuration or AI/ML CSI report configuration for SP CSI report on PUCCH.

[0161] For the MAC CE format in FIG. 3B, the LCID used for the SP CSI report on PUCCH activation/deactivation MAC CE may be reused to identify this enhanced MAC CE. The AI/ML indication field may indicate whether the followed S_i fields are used to indicate the activation / deactivation status of the SP CSI report configuration within *csi-ReportConfigToAddModList* or *AIML-CSI-ReportConfigToAddModList*. The AI/ML field is set to 1 to indicate that the S_i fields are used to indicate the activation/deactivation status of the SP CSI report configurations within *AIML-CSI-ReportConfigToAddModList*. The AI/ML field is set to 0 to indicate that the S_i fields are used to indicate the activation/deactivation status of the SP CSI report configurations within *csi-ReportConfigToAddModList*. The bitmap S_0, S_1, S_2, S_3 may indicate the

activation/deactivation status of the SP CSI report configurations within *csi-ReportConfigToAddModList* or *AIML-CSI-ReportConfigToAddModList*, S_0 refers to the report configuration which includes PUCCH resources for SP CSI reporting in the indicated BWP and has the lowest *CSI-ReportConfigId* or *AIML-CSI-ReportConfigId* within the list with type set to *semiPersistentOnPUCCH*, S_1 refers to the report configuration which includes PUCCH resources for SP CSI reporting in the indicated BWP and has the second lowest *CSI-ReportConfigId* or *AIML-CSI-ReportConfigId*, and so on. The S_i field is set to 1 to indicate that the corresponding SP CSI report configuration shall be activated. The S_i field is set to 0 to indicate that the corresponding SP CSI report configuration shall be deactivated.

[0162] However, in addition to MAC CE as shown in FIGS. 3A and 3B, a dedicated MAC CE with a dedicated LCID to activate/deactivate AI/ML CSI report on PUCCH may be introduced, as shown in FIG. 3C, and each of S_0 to S_3 may indicate whether a corresponding AI/ML CSI report configuration for SP CSI report on PUCCH is to be activated.

[0163] For MAC CE format in FIG. 3C, a dedicated LCID is used to identify this MAC CE. The bitmap S_0, S_1, S_2, S_3 may indicate the activation/deactivation status of the SP AI/ML CSI report configurations (e.g., AI/ML CSI report configurations for SP CSI report on PUCCH) within *AIML-CSI-ReportConfigToAddModList*, S_0 refers to the report configuration which includes PUCCH resources for SP AI/ML CSI reporting in the indicated BWP and has the lowest *AI/ML-CSI-ReportConfigId* within the list with type set to *semiPersistentOnPUCCH*, S_1 refers to the report configuration which includes PUCCH resources for SP AI/ML CSI reporting in the indicated BWP and has the second lowest *AI/ML-CSI-ReportConfigId*, and so on. The S_i field is set to 1 to indicate that the corresponding SP AI/ML CSI report configuration shall be activated. The S_i field is set to 0 to indicate that the corresponding SP AI/ML CSI report configuration shall be deactivated.

[0164] In some embodiments of present disclosure, the at least one CSI report configuration transmitted from the base station 102 to the UE 104 may include at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink shared channel (PUSCH), and/or at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH. In such case, for triggering a CSI report configuration for

SP CSI report on PUSCH, a DCI containing CSI request field scrambled with SP-CSI-RNTI may be reused for activation/deactivation of SP AI/ML CSI report on PUSCH.

[0165] In some embodiments of present disclosure, the UE 104 may receive, from the base station 102, a SP CSI report trigger state indication (e.g., the *CSI-SemiPersistentOnPUSCH-TriggerStateList* used in legacy CSI frame work), each CSI report trigger state (e.g., *CSI-SemiPersistentOnPUSCH-TriggerState*) indicated in the SP CSI report trigger state indication may be mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH. A CSI report trigger state indicated in the SP CSI report trigger state indication may be associated with a configuration ID of an AI/ML CSI report configuration for SP CSI report on PUSCH. For example, some of the *CSI-SemiPersistentOnPUSCH-TriggerStates*, which are mapped to a CSI request field codepoint in a DCI, may be configured with an *AIML-CSI-ReportConfigId* configured for an AI/ML CSI report configuration for SP CSI report on PUSCH, to associate an AI/ML CSI report configuration for SP CSI report on PUSCH with a CSI request field codepoint for activation. In this case, the UE may ignore the associated *CSI-ReportConfigId* for the same *CSI-SemiPersistentOnPUSCH-TriggerState* if configured.

[0166] Alternatively, in some embodiments of present disclosure, the UE 104 may receive, from the base station 102, a first SP CSI report trigger state indication (e.g., *AIML-CSI-SemiPersistentOnPUSCH-TriggerStateList*) and/or a second SP CSI report trigger state indication (e.g., the *CSI-SemiPersistentOnPUSCH-TriggerStateList*), each CSI report trigger state indicated in the first SP trigger state indication and the second SP CSI report trigger state indication (e.g. *AIML-CSI-SemiPersistentOnPUSCH-TriggerState* and/or *CSI-SemiPersistentOnPUSCH-TriggerState*) may be mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH. The first SP CSI report trigger state indication may be associated with at least one AI/ML CSI report configuration for SP CSI report on PUSCH, and the second SP CSI report trigger state indication may be associated with at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH. In other words, the UE 104 may introduce another list of *AIML-CSI-SemiPersistentOnPUSCH-TriggerState* for a serving cell, where each of *AIML-CSI-SemiPersistentOnPUSCH-TriggerState* is associated with an *AIML-CSI-ReportConfig*. Each of *AIML-CSI-SemiPersistentOnPUSCH-TriggerState* may be mapped to a CSI request field codepoint contained in a DCI scrambled with SP-

CSI-RNTI. Since both non-AI/ML CSI report and AI/ML CSI report are triggered by CSI request field contained in DCI scrambled by SP-CSI-RNTI, an additional field may be introduced in the DCI to tell the UE 104 which one or both of them are activated. For example, value 0 of the additional field may indicate non-AI/ML CSI activating; value 1 of the additional field may indicate AI/ML CSI activating, and value 2 of the additional field may indicate both non-AI/ML and AI/ML CSI activating.

[0167] In some embodiments of present disclosure, the at least one CSI report configuration transmitted from the base station 102 to the UE 104 may include at least one AI/ML CSI report configuration for aperiodic (AP) CSI report, and/or at least one non-AI/ML CSI report configuration for AP CSI report. In such case, for triggering a CSI report configuration for AP CSI report, a DCI containing a non-zero CSI request field scrambled with C-RNTI may be reused.

[0168] In some embodiments of present disclosure, the UE 104 may receive, from the base station 102, a AP CSI report trigger state indication (e.g., *CSI-AperiodicTriggerStateList* used in legacy CSI frame work), each CSI report trigger state (e.g., *CSI-AperiodicTriggerState*) indicated in the AP CSI report trigger state indication may be mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report. A CSI report trigger state indicated in the AP CSI report trigger state indication may be associated with a configuration ID of an AI/ML CSI report configuration for AP CSI report. For example, some of the *CSI-AperiodicTriggerStates* may be associated with one or more *AIML-CSI-ReportConfigs*. In this case, the UE 104 may ignore the *CSI-ReportConfig* associated with this *CSI-AperiodicTriggerState* if configured.

[0169] Alternatively, in some embodiments of present disclosure, the UE 104 may receive, from the base station 102, a first AP CSI report trigger state indication (e.g., *AIML-CSI-AperiodicTriggerStateList*) and/or a second AP CSI report trigger state indication (e.g., the *CSI-AperiodicTriggerStateList*), each CSI report trigger state indicated in the first AP CSI report trigger state indication and the second AP CSI report trigger state indication (e.g., *AIML-CSI-AperiodicTriggerState* and/or *CSI-AperiodicTriggerState*) is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report. The first AP CSI report trigger state indication may be associated with at least one AI/ML CSI report configuration for AP

CSI report, and the second AP CSI report trigger state indication may be associated with at least one non-AI/ML CSI report configuration for AP CSI report. In other words, the UE 104 may introduce another list of *AI/ML-CSI-AperiodicTriggerState* for a serving cell, where each of the *AI/ML-CSI-AperiodicTriggerState* is associated with one or more *AI/ML-CSI-ReportConfigs*. Each of the *AI/ML-CSI-AperiodicTriggerState* may be mapped with a non-zero CSI request field contained in a DCI scrambled with a C-RNTI. In order to tell the UE 104 whether a received non-zero CSI request field codepoint is used for non-AI/ML CSI report or AI/ML CSI report triggering, an additional DCI field may be introduced in the DCI. For example, value 0 of the additional DCI field may indicate non-AI/ML CSI triggering, value 1 of the additional DCI field may indicate AI/ML CSI triggering, and value 2 of the additional DCI field may indicate that both non-AI/ML and AI/ML CSI are triggered.

[0170] AI/ML based CSI including compressed CSI and predicted CSI may be basically reported as a type of UCI and may be reported by PUCCH or PUSCH. Thus, AI/ML based CSI may be multiplexed with other UCI for reporting. Similar with a legacy CSI report, there may be a case that multiple CSI reports may be reported in a same slot, but only part of them can be transmitted due to the limited UL resources. A priority rule is required to determine which CSI report should be transmitted and which one may need to be dropped.

[0171] In some embodiments of present disclosure, the UE 104 may further determine priorities among AI/ML CSI reports based on usage of the AI/ML CSI reports.

[0172] In some embodiments of present disclosure, for AI/ML CSI reports, a priority of an AI/ML CSI report for performance monitoring may be higher than a priority of an AI/ML CSI report for inference, and the priority of the AI/ML CSI report for inference may be higher than a priority of an AI/ML CSI report for data collection for model training.

[0173] In some embodiments of present disclosure, the UE 104 may further determine priorities of CSI reports (including AI/ML CSI reports and non-AI/ML CSI reports) at least based on the at least one CSI report configuration. The priorities of the CSI reports may be determined based on priority values of the CSI reports. A priority of a first CSI report is higher than a priority of a second CSI report in the case that a priority value of the first CSI report is lower than a priority value of the second CSI report.

[0174] In some embodiments of present disclosure, a priority value of a CSI report may be determined based on at least one of: a type indication value indicating a reporting type of a CSI report configuration corresponding to the CSI report, the reporting type may include one of aperiodic, SP on PUSCH, SP on PUCCH, or periodic; a configuration ID of the CSI report configuration corresponding to the CSI report; a measurement indication value indicating whether the CSI report carries L1-reference signal receiving power (RSRP) or L1-signal to interference plus noise ratio (SINR) or not; the max number of each type of CSI report configurations in the at least one CSI report configuration transmitted from the base station 102 to the UE 104, types of CSI report configuration at least include AI/ML type and non-AI/ML type; a serving cell index for the UE; and the max number of serving cells.

[0175] Determination of a priority value and a priority of a CSI report will be described in detail below. In the following description, three types of CSI report, including legacy CSI report, LTM CSI report, and AI/ML CSI report, are used as examples.

[0176] In some embodiments of present disclosure, each CSI report corresponding to a *CSI-ReportConfig*, *LTM-CSI-ReportConfig* (i.e. CSI report configuration for layer 1 and/or layer 2 (L1/L2)-triggered mobility (LTM), which may also referred to as LTM CSI report configuration) or *AIML-CSI-ReportConfig* may be associated with a priority value calculated as follows:

$$\text{Pri}_{i\text{CSI}}(y, k, c, s_1, s_2, s_3) = 2 \cdot N_{\text{cells}} \cdot M_s \cdot y + N_{\text{cells}} \cdot M_s \cdot k + M_s \cdot c + M_{s_1} + s_1, \text{ for}$$
 CSI report corresponding to *CSI-ReportConfig*.....(1)

$$\text{Pri}_{i\text{LTM-CSI}}(y, k, c, s_1, s_2, s_3) = 2 \cdot N_{\text{cells}} \cdot M_s \cdot y + N_{\text{cells}} \cdot M_s \cdot k + M_s \cdot c + s_2, \text{ for CSI}$$
 report corresponding to *LTM-CSI-ReportConfig*.....(2)

$$\text{Pri}_{i\text{AI-CSI}}(y, k, c, s_1, s_2, s_3) = 2 \cdot N_{\text{cells}} \cdot M_s \cdot y + N_{\text{cells}} \cdot M_s \cdot k + M_s \cdot c + M_{s_1} + M_{s_2} \cdot$$

$$+ s_3, \text{ for CSI report corresponding to } \textit{AIML-CSI-ReportConfig} \dots\dots(3)$$

[0177] In above equations (1) to (3), the value of y may indicate a reporting type of a CSI report, for example, $y=0$ may indicate an aperiodic CSI report(s) to be carried on PUSCH, $y=1$ may indicate a semi-persistent CSI report(s) to be carried on PUSCH, $y=2$ may indicate a semi-persistent CSI report(s) to be carried on PUCCH, and $y=3$ may indicate a periodic CSI report(s) to be carried on PUCCH. The value of k may indicate

whether the CSI report carries L1-RSRP or L1-SINR or not, for example, $k=0$ may indicate a CSI report(s) carrying L1-RSRP or L1-SINR, and $k=1$ may indicate a CSI report(s) not carrying L1-RSRP or L1-SINR. c is a serving cell index, and N_{cells} is a value of a higher layer parameter *maxNrofServingCells*, which refers to the max number of serving cells configured for the UE 104. M_s is the sum of values of the higher layer parameter *maxNrofCSI-ReportConfigurations* (indicating the max number of legacy CSI report configurations configured for the UE 104 in a serving cell), *maxNrofLTM-CSI-ReportConfigurations* (indicating the max number of LTM CSI report configurations configured for the UE 104 in a serving cell) and *maxNrofAIML-CSI-ReportConfigurations* (indicating the max number of AI/ML CSI report configurations configured for the UE 104 in a serving cell). M_{s_1} may be the value of *maxNrofLTM-CSI-ReportConfigurations*, and M_{s_2} may be the value of *maxNrofCSI-ReportConfigurations*. The value of *maxNrofCSI-ReportConfigurations*, *maxNrofLTM-CSI-ReportConfigurations* and *maxNrofAIML-CSI-ReportConfigurations* may be 0 when the corresponding CSI configuration is not configured. s_1 is *CSI-reportConfigID* of *CSI-ReportConfig* or is 0 when *CSI-ReportConfig* is not configured. s_2 is *LTM-CSI-reportConfigID* of *LTM-CSI-ReportConfig* or is 0 when *LTM-CSI-ReportConfig* is not configured. s_3 is the *AIML-CSI-reportConfigID* of *AIML-CSI-ReportConfig* or is 0 when *AIML-CSI-ReportConfig* is not configured.

[0178] Values of above parameters are only examples, other different values can be used as required. Furthermore, the difference terms in the equations (1) to (3) (for example, $(M_{s_1} + s_1)$ in equation (1), s_2 in equation (2), and $(M_{s_1} + M_{s_2} \cdot s_3)$ in equation (3)) may be used to determine a value, e.g., a value related to the configuration ID for the CSI report, to ensure that the LTM CSI report may have a higher priority than a legacy CSI report, and the legacy CSI report may have a higher priority than an AI/ML CSI report. However, equations (1) to (3) are only examples, the exact form of equations used for calculating priority values of each type of CSI reports may be designed in other ways based on the premise that the LTM CSI reports have higher priorities than legacy CSI reports, and the legacy CSI reports have higher priorities than AI/ML CSI reports.

[0179] Alternatively, in some embodiments of present disclosure, each CSI report corresponding to a *CSI-ReportConfig*, *LTM-CSI-ReportConfig* or *AIML-CSI-ReportConfig* may be associated with a priority value calculated as follows:

$$\text{Pri}_{i\text{CSI}}(y, k, c, s_1, s_2) = 2 \cdot N_{\text{cells}} \cdot M_s \cdot y + N_{\text{cells}} \cdot M_s \cdot k + M_s \cdot c + s \dots\dots(4)$$

[0180] In equation (4), the value of y may indicate a reporting type of a CSI report, for example, $y=0$ may indicate an aperiodic CSI report(s) to be carried on PUSCH, $y=1$ may indicate a semi-persistent CSI report(s) to be carried on PUSCH, $y=2$ may indicate a semi-persistent CSI report(s) to be carried on PUCCH, and $y=3$ may indicate a periodic CSI report(s) to be carried on PUCCH. The value of k may indicate whether the CSI report carries L1-RSRP or L1-SINR or not, for example, $k=0$ may indicate a CSI report(s) carrying L1-RSRP or L1-SINR, and $k=1$ may indicate a CSI report(s) not carrying L1-RSRP or L1-SINR. c is a serving cell index, and N_{cells} is a value of a higher layer parameter *maxNrofServingCells*, which refers to the max number of serving cells configured for the UE 104.

[0181] Furthermore, the M_s may be a value of a higher layer parameter *maxNrofCSI-ReportConfigurations* for CSI corresponding to *CSI-ReportConfig*, or *maxNrofLTM-CSI-ReportConfigurations* for CSI corresponding to *LTM-CSI-ReportConfig*, or *maxNrofAIML-CSI-ReportConfigurations* for CSI corresponding to *AIML-CSI-ReportConfig*. s is *CSI-reportConfigID* of *CSI-ReportConfig*, or *LTM-CSI-reportConfigID* of *LTM-CSI-ReportConfig*, or *AIML-CSI-reportConfigID* of *AIML-CSI-ReportConfig*.

[0182] In some embodiments of present disclosure, in the case that time occupancy of PUSCH or PUCCH to carry CSI reports overlap in at least one orthogonal frequency-division multiplexing (OFDM) symbol and the CSI reports are transmitted on the same carrier, the UE 104 may determine reporting priorities of the CSI reports based on a predetermined rule, and drop one of the CSI reports based on the reporting priorities.

[0183] In some embodiments of present disclosure, the predetermined rule may be: an AI/ML CSI report has a lower reporting priority than a non-AI/ML CSI report.

[0184] In some other embodiments of present disclosure, the predetermined rule may include allocating reporting priorities, in a priority order from high to low, to an AP non-AI/ML CSI report, an AP AI/ML CSI report, a SP non-AI/ML CSI report on PUSCH, a SP AI/ML CSI report on PUSCH, a SP non-AI/ML CSI report on PUCCH, a SP AI/ML CSI report on PUCCH, a periodic non-AI/ML CSI report, and a periodic AI/ML CSI report. For example, the priority order may be:

AP LTM CSI report > AP Legacy CSI report > AP AI/ML CSI report > SP LTM CSI report on PUSCH > SP legacy CSI report on PUSCH > SP AI/ML CSI report on PUSCH > SP LTM CSI report on PUCCH > SP legacy CSI report on PUCCH > SP AI/ML CSI report on PUCCH > Periodic LTM CSI report > Periodic legacy CSI report > Periodic AI/ML CSI report.

[0185] In some other embodiments of present disclosure, the predetermined rule may include allocating reporting priorities, in a priority order from high to low, to an AP non-AI/ML CSI report, a SP non-AI/ML CSI report on PUSCH, a SP non-AI/ML CSI report on PUCCH, a periodic non-AI/ML CSI report, an AP AI/ML CSI report, a SP AI/ML CSI report on PUSCH, a SP AI/ML CSI report on PUCCH, and a periodic AI/ML CSI report. For example, the priority order may be:

AP LTM CSI report > SP LTM CSI report on PUSCH > SP LTM CSI report on PUCCH > Periodic LTM CSI report > AP Legacy CSI report > SP legacy CSI report on PUSCH > SP legacy CSI report on PUCCH > Periodic legacy CSI report > AP AI/ML CSI report > SP AI/ML CSI report on PUSCH > SP AI/ML CSI report on PUCCH > Periodic AI/ML CSI report.

[0186] FIG. 4 illustrates an example of a device 400 that supports CSI reporting in accordance with aspects of the present disclosure. The device 400 may be an example of a UE 104 as described herein. The device 400 may support wireless communication with one or more network entities 102, UEs 104, or any combination thereof. The device 400 may include components for bi-directional communications including components for transmitting and receiving communications, such as a processor 402, a memory 404, a transceiver 406, and, optionally, an I/O controller 408. These components may be in electronic communication or otherwise coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0187] The processor 402, the memory 404, the transceiver 406, or various combinations thereof or various components thereof may be examples of means for performing various aspects of the present disclosure as described herein. For example, the processor 402, the memory 404, the transceiver 406, or various combinations or components thereof may support a method for performing one or more of the operations described herein.

[0188] In some implementations, the processor 402, the memory 404, the transceiver 406, or various combinations or components thereof may be implemented in hardware (e.g., in communications management circuitry). The hardware may include a processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), a field-programmable gate array (FPGA) or other programmable logic device, a discrete gate or transistor logic, discrete hardware components, or any combination thereof configured as or otherwise supporting a means for performing the functions described in the present disclosure. In some implementations, the processor 402 and the memory 404 coupled with the processor 402 may be configured to perform one or more of the functions described herein (e.g., executing, by the processor 402, instructions stored in the memory 404).

[0189] For example, the processor 402 may support wireless communication at the device 400 in accordance with examples as disclosed herein. The processor 402 may be configured to operable to support a means for transmitting, via a transceiver 406 and to a base station 102, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and means for receiving, via the transceiver 406 and from the base station 102, at least one CSI report configuration, the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0190] The processor 402 may include an intelligent hardware device (e.g., a general-purpose processor, a DSP, a CPU, a microcontroller, an ASIC, an FPGA, a programmable logic device, a discrete gate or transistor logic component, a discrete hardware component, or any combination thereof). In some implementations, the processor 402 may be configured to operate a memory array using a memory controller. In some other implementations, a memory controller may be integrated into the processor 402. The processor 402 may be configured to execute computer-readable instructions stored in a memory (e.g., the memory 404) to cause the device 400 to perform various functions of the present disclosure.

[0191] The memory 404 may include random access memory (RAM) and read-only memory (ROM). The memory 404 may store computer-readable, computer-executable code including instructions that, when executed by the processor 402 cause the device 400 to perform various functions described herein. The code may be stored in a non-

transitory computer-readable medium such as system memory or another type of memory. In some implementations, the code may not be directly executable by the processor 402 but may cause a computer (e.g., when compiled and executed) to perform functions described herein. In some implementations, the memory 404 may include, among other things, a basic I/O system (BIOS) which may control basic hardware or software operation such as the interaction with peripheral components or devices.

[0192] The I/O controller 408 may manage input and output signals for the device 400. The I/O controller 408 may also manage peripherals not integrated into the device 400. In some implementations, the I/O controller 408 may represent a physical connection or port to an external peripheral. In some implementations, the I/O controller 408 may utilize an operating system such as iOS®, ANDROID®, MS-WINDOWS®, OS/2®, UNIX®, LINUX®, or another known operating system. In some implementations, the I/O controller 408 may be implemented as part of a processor, such as the processor 406. In some implementations, a user may interact with the device 400 via the I/O controller 408 or via hardware components controlled by the I/O controller 408.

[0193] In some implementations, the device 400 may include a single antenna 410. However, in some other implementations, the device 400 may have more than one antenna 410 (i.e., multiple antennas), including multiple antenna panels or antenna arrays, which may be capable of concurrently transmitting or receiving multiple wireless transmissions. The transceiver 406 may communicate bi-directionally, via the one or more antennas 410, wired, or wireless links as described herein. For example, the transceiver 406 may represent a wireless transceiver and may communicate bi-directionally with another wireless transceiver. The transceiver 406 may also include a modem to modulate the packets, to provide the modulated packets to one or more antennas 410 for transmission, and to demodulate packets received from the one or more antennas 410. The transceiver 406 may include one or more transmit chains, one or more receive chains, or a combination thereof.

[0194] A transmit chain may be configured to generate and transmit signals (e.g., control information, data, packets). The transmit chain may include at least one modulator for modulating data onto a carrier signal, preparing the signal for transmission over a wireless medium. The at least one modulator may be configured to support one or more

techniques such as amplitude modulation (AM), frequency modulation (FM), or digital modulation schemes like phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmit chain may also include at least one power amplifier configured to amplify the modulated signal to an appropriate power level suitable for transmission over the wireless medium. The transmit chain may also include one or more antennas 610 for transmitting the amplified signal into the air or wireless medium.

[0195] A receive chain may be configured to receive signals (e.g., control information, data, packets) over a wireless medium. For example, the receive chain may include one or more antennas 610 for receive the signal over the air or wireless medium. The receive chain may include at least one amplifier (e.g., a low-noise amplifier (LNA)) configured to amplify the received signal. The receive chain may include at least one demodulator configured to demodulate the receive signal and obtain the transmitted data by reversing the modulation technique applied during transmission of the signal. The receive chain may include at least one decoder for decoding the processing the demodulated signal to receive the transmitted data.

[0196] FIG. 5 illustrates an example of a device 500 that supports CSI reporting in accordance with aspects of the present disclosure. The device 500 may be an example of a base station 104 as described herein, the base station 102 may include a plurality of TRPs. The device 500 may support wireless communication with one or more network entities 102, UEs 104, or any combination thereof. The device 500 may include components for bi-directional communications including components for transmitting and receiving communications, such as a processor 502, a memory 504, a transceiver 506, and, optionally, an I/O controller 508. These components may be in electronic communication or otherwise coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0197] The processor 502, the memory 504, the transceiver 506, or various combinations thereof or various components thereof may be examples of means for performing various aspects of the present disclosure as described herein. For example, the processor 502, the memory 504, the transceiver 506, or various combinations or components thereof may support a method for performing one or more of the operations described herein.

[0198] In some implementations, the processor 502, the memory 504, the transceiver 506, or various combinations or components thereof may be implemented in hardware (e.g., in communications management circuitry). The hardware may include a processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), a field-programmable gate array (FPGA) or other programmable logic device, a discrete gate or transistor logic, discrete hardware components, or any combination thereof configured as or otherwise supporting a means for performing the functions described in the present disclosure. In some implementations, the processor 502 and the memory 504 coupled with the processor 502 may be configured to perform one or more of the functions described herein (e.g., executing, by the processor 502, instructions stored in the memory 504).

[0199] For example, the processor 502 may support wireless communication at the device 500 in accordance with examples as disclosed herein. The processor 502 may be configured to operable to support a means for receiving, via the transceiver 506 and from a user equipment (UE) 104, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and means for transmitting, via the transceiver 506 and to the UE 104, at least one CSI report configuration, the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0200] The processor 502 may include an intelligent hardware device (e.g., a general-purpose processor, a DSP, a CPU, a microcontroller, an ASIC, an FPGA, a programmable logic device, a discrete gate or transistor logic component, a discrete hardware component, or any combination thereof). In some implementations, the processor 502 may be configured to operate a memory array using a memory controller. In some other implementations, a memory controller may be integrated into the processor 502. The processor 502 may be configured to execute computer-readable instructions stored in a memory (e.g., the memory 504) to cause the device 500 to perform various functions of the present disclosure.

[0201] The memory 504 may include random access memory (RAM) and read-only memory (ROM). The memory 504 may store computer-readable, computer-executable code including instructions that, when executed by the processor 502 cause the device 500 to perform various functions described herein. The code may be stored in a non-

transitory computer-readable medium such as system memory or another type of memory. In some implementations, the code may not be directly executable by the processor 502 but may cause a computer (e.g., when compiled and executed) to perform functions described herein. In some implementations, the memory 504 may include, among other things, a basic I/O system (BIOS) which may control basic hardware or software operation such as the interaction with peripheral components or devices.

[0202] The I/O controller 508 may manage input and output signals for the device 500. The I/O controller 508 may also manage peripherals not integrated into the device 500. In some implementations, the I/O controller 508 may represent a physical connection or port to an external peripheral. In some implementations, the I/O controller 508 may utilize an operating system such as iOS®, ANDROID®, MS-WINDOWS®, OS/2®, UNIX®, LINUX®, or another known operating system. In some implementations, the I/O controller 508 may be implemented as part of a processor, such as the processor 506. In some implementations, a user may interact with the device 500 via the I/O controller 508 or via hardware components controlled by the I/O controller 508.

[0203] In some implementations, the device 500 may include a single antenna 510. However, in some other implementations, the device 500 may have more than one antenna 510 (i.e., multiple antennas), including multiple antenna panels or antenna arrays, which may be capable of concurrently transmitting or receiving multiple wireless transmissions. The transceiver 506 may communicate bi-directionally, via the one or more antennas 510, wired, or wireless links as described herein. For example, the transceiver 506 may represent a wireless transceiver and may communicate bi-directionally with another wireless transceiver. The transceiver 506 may also include a modem to modulate the packets, to provide the modulated packets to one or more antennas 510 for transmission, and to demodulate packets received from the one or more antennas 510. The transceiver 506 may include one or more transmit chains, one or more receive chains, or a combination thereof.

[0204] A transmit chain may be configured to generate and transmit signals (e.g., control information, data, packets). The transmit chain may include at least one modulator for modulating data onto a carrier signal, preparing the signal for transmission over a wireless medium. The at least one modulator may be configured to support one or more

techniques such as amplitude modulation (AM), frequency modulation (FM), or digital modulation schemes like phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmit chain may also include at least one power amplifier configured to amplify the modulated signal to an appropriate power level suitable for transmission over the wireless medium. The transmit chain may also include one or more antennas 710 for transmitting the amplified signal into the air or wireless medium.

[0205] A receive chain may be configured to receive signals (e.g., control information, data, packets) over a wireless medium. For example, the receive chain may include one or more antennas 510 for receive the signal over the air or wireless medium. The receive chain may include at least one amplifier (e.g., a low-noise amplifier (LNA)) configured to amplify the received signal. The receive chain may include at least one demodulator configured to demodulate the receive signal and obtain the transmitted data by reversing the modulation technique applied during transmission of the signal. The receive chain may include at least one decoder for decoding the processing the demodulated signal to receive the transmitted data.

[0206] FIG. 6 illustrates an example of a processor 600 that supports CSI reporting in accordance with aspects of the present disclosure. The processor 600 may be an example of a processor configured to perform various operations in accordance with examples as described herein. The processor 600 may include a controller 602 configured to perform various operations in accordance with examples as described herein. The processor 600 may optionally include at least one memory 604, such as L1/L2/L3 cache. Additionally, or alternatively, the processor 600 may optionally include one or more arithmetic-logic units (ALUs) 600. One or more of these components may be in electronic communication or otherwise coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0207] The processor 600 may be a processor chipset and include a protocol stack (e.g., a software stack) executed by the processor chipset to perform various operations (e.g., receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) in accordance with examples as described herein. The processor chipset may include one or more cores, one or more caches (e.g., memory local to or included in the processor chipset (e.g., the processor 600)) or other memory (e.g., random access memory (RAM), read-only memory (ROM),

dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), static RAM (SRAM), ferroelectric RAM (FeRAM), magnetic RAM (MRAM), resistive RAM (RRAM), flash memory, phase change memory (PCM), and others).

[0208] The controller 602 may be configured to manage and coordinate various operations (e.g., signaling, receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) of the processor 600 to cause the processor 600 to support various operations of a base station in accordance with examples as described herein. For example, the controller 602 may operate as a control unit of the processor 600, generating control signals that manage the operation of various components of the processor 600. These control signals include enabling or disabling functional units, selecting data paths, initiating memory access, and coordinating timing of operations.

[0209] The controller 602 may be configured to fetch (e.g., obtain, retrieve, receive) instructions from the memory 604 and determine subsequent instruction(s) to be executed to cause the processor 600 to support various operations in accordance with examples as described herein. The controller 602 may be configured to track memory address of instructions associated with the memory 604. The controller 602 may be configured to decode instructions to determine the operation to be performed and the operands involved. For example, the controller 602 may be configured to interpret the instruction and determine control signals to be output to other components of the processor 600 to cause the processor 600 to support various operations in accordance with examples as described herein. Additionally, or alternatively, the controller 602 may be configured to manage flow of data within the processor 600. The controller 602 may be configured to control transfer of data between registers, arithmetic logic units (ALUs), and other functional units of the processor 600.

[0210] The memory 604 may include one or more caches (e.g., memory local to or included in the processor 600 or other memory, such RAM, ROM, DRAM, SDRAM, SRAM, MRAM, flash memory, etc. In some implementation, the memory 604 may reside within or on a processor chipset (e.g., local to the processor 600). In some other implementations, the memory 604 may reside external to the processor chipset (e.g., remote to the processor 600).

[0211] The memory 604 may store computer-readable, computer-executable code including instructions that, when executed by the processor 600, cause the processor 600 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as system memory or another type of memory. The controller 602 and/or the processor 600 may be configured to execute computer-readable instructions stored in the memory 604 to cause the processor 600 to perform various functions. For example, the processor 600 and/or the controller 602 may be coupled with or to the memory 604, and the processor 600, the controller 602, and the memory 604 may be configured to perform various functions described herein. In some examples, the processor 600 may include multiple processors and the memory 604 may include multiple memories. One or more of the multiple processors may be coupled with one or more of the multiple memories, which may, individually or collectively, be configured to perform various functions herein.

[0212] The one or more ALUs 600 may be configured to support various operations in accordance with examples as described herein. In some implementation, the one or more ALUs 600 may reside within or on a processor chipset (e.g., the processor 600). In some other implementations, the one or more ALUs 600 may reside external to the processor chipset (e.g., the processor 600). One or more ALUs 600 may perform one or more computations such as addition, subtraction, multiplication, and division on data. For example, one or more ALUs 600 may receive input operands and an operation code, which determines an operation to be executed. One or more ALUs 600 be configured with a variety of logical and arithmetic circuits, including adders, subtractors, shifters, and logic gates, to process and manipulate the data according to the operation. Additionally, or alternatively, the one or more ALUs 600 may support logical operations such as AND, OR, exclusive-OR (XOR), not-OR (NOR), and not-AND (NAND), enabling the one or more ALUs 600 to handle conditional operations, comparisons, and bitwise operations.

[0213] The processor 600 may support wireless communication in accordance with examples as disclosed herein. The processor 600 may be configured to or operable to support a means for transmitting, via the transceiver and to a base station, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and means for receiving, via the transceiver and from the base station, at least one CSI report configuration, the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML

functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0214] FIG. 7 illustrates an example of a processor 700 that supports CSI reporting in accordance with aspects of the present disclosure. The processor 700 may be an example of a processor configured to perform various operations in accordance with examples as described herein. The processor 700 may include a controller 702 configured to perform various operations in accordance with examples as described herein. The processor 700 may optionally include at least one memory 704, such as L1/L2/L3 cache. Additionally, or alternatively, the processor 700 may optionally include one or more arithmetic-logic units (ALUs) 700. One or more of these components may be in electronic communication or otherwise coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0215] The processor 700 may be a processor chipset and include a protocol stack (e.g., a software stack) executed by the processor chipset to perform various operations (e.g., receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) in accordance with examples as described herein. The processor chipset may include one or more cores, one or more caches (e.g., memory local to or included in the processor chipset (e.g., the processor 700) or other memory (e.g., random access memory (RAM), read-only memory (ROM), dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), static RAM (SRAM), ferroelectric RAM (FeRAM), magnetic RAM (MRAM), resistive RAM (RRAM), flash memory, phase change memory (PCM), and others).

[0216] The controller 702 may be configured to manage and coordinate various operations (e.g., signaling, receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) of the processor 700 to cause the processor 700 to support various operations of a UE in accordance with examples as described herein. For example, the controller 702 may operate as a control unit of the processor 700, generating control signals that manage the operation of various components of the processor 700. These control signals include enabling or disabling functional units, selecting data paths, initiating memory access, and coordinating timing of operations.

[0217] The controller 702 may be configured to fetch (e.g., obtain, retrieve, receive) instructions from the memory 704 and determine subsequent instruction(s) to be executed to cause the processor 700 to support various operations in accordance with examples as described herein. The controller 702 may be configured to track memory address of instructions associated with the memory 704. The controller 702 may be configured to decode instructions to determine the operation to be performed and the operands involved. For example, the controller 702 may be configured to interpret the instruction and determine control signals to be output to other components of the processor 700 to cause the processor 700 to support various operations in accordance with examples as described herein. Additionally, or alternatively, the controller 702 may be configured to manage flow of data within the processor 700. The controller 702 may be configured to control transfer of data between registers, arithmetic logic units (ALUs), and other functional units of the processor 700.

[0218] The memory 704 may include one or more caches (e.g., memory local to or included in the processor 700 or other memory, such RAM, ROM, DRAM, SDRAM, SRAM, MRAM, flash memory, etc. In some implementation, the memory 704 may reside within or on a processor chipset (e.g., local to the processor 700). In some other implementations, the memory 704 may reside external to the processor chipset (e.g., remote to the processor 700).

[0219] The memory 704 may store computer-readable, computer-executable code including instructions that, when executed by the processor 700, cause the processor 700 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as system memory or another type of memory. The controller 702 and/or the processor 700 may be configured to execute computer-readable instructions stored in the memory 704 to cause the processor 700 to perform various functions. For example, the processor 700 and/or the controller 702 may be coupled with or to the memory 704, and the processor 700, the controller 702, and the memory 704 may be configured to perform various functions described herein. In some examples, the processor 700 may include multiple processors and the memory 704 may include multiple memories. One or more of the multiple processors may be coupled with one or more of the multiple memories, which may, individually or collectively, be configured to perform various functions herein.

[0220] The one or more ALUs 700 may be configured to support various operations in accordance with examples as described herein. In some implementation, the one or more ALUs 700 may reside within or on a processor chipset (e.g., the processor 700). In some other implementations, the one or more ALUs 700 may reside external to the processor chipset (e.g., the processor 700). One or more ALUs 700 may perform one or more computations such as addition, subtraction, multiplication, and division on data. For example, one or more ALUs 700 may receive input operands and an operation code, which determines an operation to be executed. One or more ALUs 700 be configured with a variety of logical and arithmetic circuits, including adders, subtractors, shifters, and logic gates, to process and manipulate the data according to the operation. Additionally, or alternatively, the one or more ALUs 700 may support logical operations such as AND, OR, exclusive-OR (XOR), not-OR (NOR), and not-AND (NAND), enabling the one or more ALUs 700 to handle conditional operations, comparisons, and bitwise operations.

[0221] The processor 700 may support wireless communication in accordance with examples as disclosed herein. The processor 700 may be configured to or operable to support means for receiving, via the transceiver and from a user equipment (UE), capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities; and means for transmitting, via the transceiver and to the UE, at least one CSI report configuration, the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.

[0222] FIG. 8 illustrates a flowchart of a method 800 that supports CSI reporting in accordance with aspects of the present disclosure. The operations of the method 900 may be implemented by a device or its components as described herein. For example, the operations of the method 800 may be performed by the UE 104 as described herein. In some implementations, the device may execute a set of instructions to control the function elements of the device to perform the described functions. Additionally, or alternatively, the device may perform aspects of the described functions using special-purpose hardware.

[0223] At 805, the method may include transmitting, via the transceiver and to a base station, capability information on support of one or more artificial intelligence or machine

learning (AI/ML) functionalities. The operations of 805 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 805 may be performed by a device as described with reference to FIG. 1.

[0224] At 810, the method may include receiving, via the transceiver and from the base station, at least one CSI report configuration, the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models. The operations of 810 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 810 may be performed by a device as described with reference to FIG. 1.

[0225] FIG. 9 illustrates a flowchart of a method 900 that supports CSI reporting in accordance with aspects of the present disclosure. The operations of the method 900 may be implemented by a device or its components as described herein. For example, the operations of the method 900 may be performed by the base station 102 as described herein. In some implementations, the device may execute a set of instructions to control the function elements of the device to perform the described functions. Additionally, or alternatively, the device may perform aspects of the described functions using special-purpose hardware.

[0226] At 905, the method may include receiving, via the transceiver and from a user equipment (UE), capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities. The operations of 905 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 905 may be performed by a device as described with reference to FIG. 1.

[0227] At 910, the method may include transmitting, via the transceiver and to the UE, at least one CSI report configuration, wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models. The operations of 910 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 910 may be performed by a device as described with reference to FIG. 1.

[0228] It should be noted that the methods described herein describes possible implementations, and that the operations and the steps may be rearranged or otherwise modified and that other implementations are possible. Further, aspects from two or more of the methods may be combined.

[0229] The various illustrative blocks and components described in connection with the disclosure herein may be implemented or performed with a general-purpose processor, a DSP, an ASIC, a CPU, an FPGA or other programmable logic device, discrete gate or transistor logic, discrete hardware components, or any combination thereof designed to perform the functions described herein. A general-purpose processor may be a microprocessor, but in the alternative, the processor may be any processor, controller, microcontroller, or state machine. A processor may also be implemented as a combination of computing devices (e.g., a combination of a DSP and a microprocessor, multiple microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

[0230] The functions described herein may be implemented in hardware, software executed by a processor, firmware, or any combination thereof. If implemented in software executed by a processor, the functions may be stored on or transmitted over as one or more instructions or code on a computer-readable medium. Other examples and implementations are within the scope of the disclosure and appended claims. For example, due to the nature of software, functions described herein may be implemented using software executed by a processor, hardware, firmware, hardwiring, or combinations of any of these. Features implementing functions may also be physically located at various positions, including being distributed such that portions of functions are implemented at different physical locations.

[0231] Computer-readable media includes both non-transitory computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A non-transitory storage medium may be any available medium that may be accessed by a general-purpose or special-purpose computer. By way of example, non-transitory computer-readable media may include RAM, ROM, electrically erasable programmable ROM (EEPROM), flash memory, compact disk (CD) ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other non-transitory medium that may be used to carry

or store desired program code means in the form of instructions or data structures and that may be accessed by a general-purpose or special-purpose computer, or a general-purpose or special-purpose processor.

[0232] As used herein, including in the claims, an article “a” before an element is unrestricted and understood to refer to “at least one” of those elements or “one or more” of those elements. The terms “a,” “at least one,” “one or more,” and “at least one of one or more” may be interchangeable. As used herein, including in the claims, “or” as used in a list of items (e.g., a list of items prefaced by a phrase such as “at least one of” or “one or more of” or “one or both of”) indicates an inclusive list such that, for example, a list of at least one of A, B, or C means A or B or C or AB or AC or BC or ABC (i.e., A and B and C). Also, as used herein, the phrase “based on” shall not be construed as a reference to a closed set of conditions. For example, an example step that is described as “based on condition A” may be based on both a condition A and a condition B without departing from the scope of the present disclosure. In other words, as used herein, the phrase “based on” shall be construed in the same manner as the phrase “based at least in part on. Further, as used herein, including in the claims, a “set” may include one or more elements.

[0233] The description herein is provided to enable a person having ordinary skill in the art to make or use the disclosure. Various modifications to the disclosure will be apparent to a person having ordinary skill in the art, and the generic principles defined herein may be applied to other variations without departing from the scope of the disclosure. Thus, the disclosure is not limited to the examples and designs described herein but is to be accorded the broadest scope consistent with the principles and novel features disclosed herein.

CLAIMS

What is claimed is:

1. A user equipment (UE), comprising:
a processor; and
a transceiver coupled to the processor,
wherein the processor is configured to:
transmit, via the transceiver and to a base station, capability information on support of one or more artificial intelligence or machine learning (AI/ML) functionalities;
and
receive, via the transceiver and from the base station, at least one CSI report configuration,
wherein the at least one CSI report configuration at least includes at least one AI/ML CSI report configuration associated with at least one AI/ML functionality or at least one AI/ML model, and an AI/ML functionality is associated with one or more AI/ML models.
2. The UE of claim 1, wherein an AI/ML CSI report configuration among the at least one AI/ML CSI report configuration includes at least one of:
a configuration identifier (ID) indicating the AI/ML CSI report configuration identity;
a resource information on CSI resource for channel measurement or CSI resource for interference measurement;
an AI/ML functionality or model information indicating an AI/ML functionality or model associated with the AI/ML CSI report configuration;
a usage information indicating the usage of the AI/ML CSI report configuration;
a reporting type for the AI/ML CSI report configuration, wherein the reporting type includes one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic;
a rank subset restriction indicating an allowed rank indicator (RI) report for the AI/ML CSI report configuration; and
an associated CSI report configuration for fallback operation, which indicates a non-AI/ML CSI report configuration for fallback operation without AI/ML operation.

3. The UE of claim 2, wherein in the case that the UE determines to perform a fallback operation without AI/ML operation based on the AI/ML CSI report configuration, the processor is configured to:

drop an AI/ML CSI report corresponding to the AI/ML CSI report configuration, and transmit, via the transceiver and to the base station, a non-AI/ML CSI report corresponding to the associated non-AI/ML CSI report configuration, wherein the same CSI resources for channel and/or interference measurement for the AI/ML CSI report configuration is used for the associated non-AI/ML CSI report configuration.

4. The UE of claim 2, wherein

in case that there are more than one AI/ML functionality identified at the UE, the AI/ML functionality or model information includes a functionality ID of the AI/ML functionality associated with the AI/ML CSI report configuration; and

in case that there is only one AI/ML functionality identified at the UE, the AI/ML functionality or model information includes an AI/ML model ID for the AI/ML functionality associated with the AI/ML CSI report configuration, or the AI/ML functionality or model information indicates to enable the AI/ML functionality associated with the AI/ML CSI report configuration.

5. The UE of claim 2, wherein the usage of the AI/ML CSI report configuration includes at least one of: data collection for model training, model inference, and performance monitoring,

wherein the model inference includes at least one of model inference for CSI compression, model inference for CSI prediction, model inference for spatial domain beam prediction, and model inference for temporal domain beam prediction.

6. The UE of claim 5, wherein in the case that the AI/ML CSI report configuration is configured for data collection for model training for two-sided model for CSI compression, a first CSI obtained by using an AI/ML model at the UE together with a second CSI obtained without using any AI/ML model at the UE, which is used as a model input for the AI/ML model at the UE, are included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

7. The UE of claim 5, wherein in the case that the AI/ML CSI report configuration is configured for inference for CSI compression, and a number of uplink control information (UCI) bits for a compressed CSI report is determined by the UE,

the number of UCI bits for the compressed CSI report is included in a first part of an AI/ML CSI report configured based on the AI/ML CSI report configuration, and a compressed CSI is included in a second part of the AI/ML CSI report.

8. The UE of claim 5, wherein in the case that the AI/ML CSI report configuration is configured for performance monitoring,

a performance gap between a reported CSI and the actual CSI of a same CSI-RS resource on a same transmission occasion is included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

9. The UE of claim 5, wherein in the case that the AI/ML CSI report configuration is configured for performance monitoring,

a ground-truth CSI corresponding to a CSI-RS resource or a CSI-RS resource set with a certain quantization scheme is included in an AI/ML CSI report configured based on the AI/ML CSI report configuration.

10. The UE of claim 1, wherein the at least one CSI report configuration includes at least one of:

at least one AI/ML CSI report configuration for semi-persistent (SP) CSI report on physical uplink shared channel (PUSCH), and

at least one non-AI/ML CSI report configuration for SP CSI report on PUSCH.

11. The UE of claim 10, wherein the processor is further configured to:

receive, via the transceiver and from the base station, a SP CSI report trigger state indication, each CSI report trigger state indicated in the SP CSI report trigger state indication is mapped to a CSI request field codepoint contained in a DCI format used for triggering a SP CSI report on PUSCH,

wherein a CSI report trigger state indicated in the SP CSI report trigger state indication is associated with a configuration identifier (ID) of an AI/ML CSI report configuration for SP CSI report on PUSCH.

12. The UE of claim 1, wherein the at least one CSI report configuration includes at least one of:

at least one AI/ML CSI report configuration for aperiodic (AP) CSI report, and
at least one non-AI/ML CSI report configuration for AP CSI report.

13. The UE of claim 12, wherein the processor is further configured to:

receive, via the transceiver and from the base station, a AP CSI report trigger state indication, each CSI report trigger state indicated in the AP CSI report trigger state indication is mapped to a non-zero CSI request field codepoint in a DCI format used for triggering an AP CSI report,

wherein a CSI report trigger state indicated in the AP CSI report trigger state indication is associated with a configuration ID of an AI/ML CSI report configuration for AP CSI report.

14. The UE of claim 5, wherein a priority of an AI/ML CSI report for performance monitoring is higher than a priority of an AI/ML CSI report for inference, and the priority of the AI/ML CSI report for inference is higher than a priority of an AI/ML CSI report for data collection for model training.

15. The UE of claim 1, wherein priorities of the CSI reports are determined based on priority values of the CSI reports,

wherein a priority value of a CSI report is determined based on at least one of :

a type indication value indicating a reporting type of a CSI report configuration corresponding to the CSI report, wherein the reporting type includes one of aperiodic, semi-persistent (SP) on physical uplink shared channel (PUSCH), SP on physical uplink control channel (PUCCH), or periodic;

a configuration identifier (ID) of the CSI report configuration corresponding to the CSI report;

a measurement indication value indicating whether the CSI report carries L1-reference signal receiving power (RSRP) or L1-signal to interference plus noise ratio (SINR) or not;

the max number of each type of CSI report configurations in the at least one CSI report configuration, wherein types of CSI report configuration at least include AI/ML type and non-AI/ML type;

a serving cell index for the UE; and
the max number of serving cells.

16. The UE of claim 1, wherein predetermined rule includes: allocating reporting priorities, in a priority order from high to low, to an AP non-AI/ML CSI report, a SP non-AI/ML CSI report on PUSCH, a SP non-AI/ML CSI report on PUCCH, a periodic non-AI/ML CSI report, an AP AI/ML CSI report, a SP AI/ML CSI report on PUSCH, a SP AI/ML CSI report on PUCCH, and a periodic AI/ML CSI report.